Defining HLA-A locus alleles from DNA using ARMS-PCR.

Thesis

UNRESTRICTED

Defining HLA-A locus alleles from DNA using ARMS-PCR.

A thesis submitted by Peter Krausa B.Sc.

IN PARTIAL FULFILMENT
OF THE REQUIREMENTS
OF THE OPEN UNIVERSITY
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY
January 1997

Sponsoring Establishment;
Imperial Cancer Research Fund,
Institute of Molecular Medicine,
Oxford, United Kingdom.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

Dedicated to Tracey Krausa.

Acknowledgments.

During the course of this research, I have had the privilege to work with a number of people, both in Oxford and further afield. Amongst this group, I am grateful to my supervisors for the help offered during the course of this work. Particularly to Andrew McMichael for support, and advice and Michael Browning, who persuaded me that a Ph.D would be a good idea. Many thanks to the I.C.R.F., for allowing me to pursue this degree. Thanks also to Applied Biosystems for giving me their time to complete writing. I would also like to thank members of both of the laboratories I have been lucky enough to work with, at the Institute of Molecular Medicine. Many thanks for your friendship and support.

The project has also allowed me to collaborate with a number of laboratories, who have provided valuable samples, or expertise to parts of this work. Samples were provided by David Savage and members of the Singapore Blood Transfusion Service, by Adrian Hill, Carlo Carcassi, Kelly McDonald, Mike Bunce, Frances Gotch and Steve McAdam. I would also like to thank members of Professor Rammensee's Laboratory in Tübingen, Germany namely Stefan Stefanovic and particularly Christian Münz for persevering and squeezing out the peptide data which finally tied things up.

On a personal level, I would like to thank those whose friendship has been more than important to me in my time in Oxford. Particularly to those at 6 Lime Walk for being such fun and whose friendship I really value. Thanks to Steve McAdam for a multitude of favours and not knowing how to say no, I appreciate it. Also to Mike and Vickie Browning whose friendship has been important, both in the good and not so good times of this Ph.D. And to Tracey Krausa for tremendous encouragement (and tolerance) throughout the whole of the last few years. I know you'll be glad to see this submitted.

The last words go to Mike Browning. As mentioned, he suggested I consider pursuing a Ph.D., for which I am really grateful. He has been important to all aspects of this work, from supervision to moral support. His comments and suggestions have always proved more than useful, as has his friendship over the last few years. Cheers to you all.

Abstract.

Defining HLA-A locus Alleles from DNA using ARMS-PCR.

Submitted for the degree of Doctor of Philosophy by Peter Krausa. B.Sc.

The HLA system is a highly polymorphic group of genes which play a central role in the immune response. The definition of the HLA system makes an important contribution to a number of fields. In the clinical setting it is a requirement for transplantation. For anthropology, the HLA system is a useful marker for assessing population groups. Functionally, it is an important component of the T cell mediated immune response. The definition offered by the established method of serology to HLA class I tissue-typing is restrictive since serology fails to discriminate between alleles which are functionally distinct. This research represents one of the first DNA based approaches and the first by PCR-SSP, to define HLA class I specificities, with particularly reference to the HLA-A locus. This thesis charts both the development and application of a robust, simple to perform, easy to interpret yet highly powerful PCR based system in which allelic definition is demonstrated. In addition to resolution, a DNA approach offers many practical advantages over the restrictive demands of a serological typing. Further more, this thesis examines the functional aspect of polymorphism, in defining the peptide binding preferences of a group of closely related HLA specificities. By so doing, this thesis shows the relevance of a highly definitive approach to HLA typing.

Defining HLA-A locus alleles from DNA using ARMS-PCR.

Table of contents

Illustration list.

Table List.

Abbreviations.

ARMS	Amplification Refractory Mutation System	MHC	Major Histocompatibility complex
β2m	Beta-2-microglobulin	MLC	Mixed Lymphocyte culture
BCL	B cell line	MLR	Mixed Lymphocyte reaction
CTL	Cytotoxic T Lymphocyte	NK	Natural Killer cell
CREG	Cross Reactive group	PBMC	Peripheral Blood Mononuclear Cells
EBV	Epstein Barr Virus	PCR	Polymerase Chain Reaction
FACS	Fluorescence Activated Cell Sorting	RFLP	Restriction Fragment Length Polymorphism
GVHD	Graft versus Host Disease	RT-PCR	Reverse Transcriptase - Polymerase Chain Reaction
HIFCS	Heat Inactivated Foetal Calf Serum	SBT	Sequence Based Typing
HIV	Human Immunodeficiency Virus	SSCP	Single Stranded Conformational Polymorphism
HLA	Human Leukocyte A or Antigen	SSO(P)	Sequence Specific Oligonucleotide (Probe)
HPLC	High Pressure Liquid Chromatography	SSP	Sequence Specific Primer
IEF	Iso-Electric Focusing	TAP	Transporter Associated with Antigen processing
IFN	Interferon	TCR	T Cell Receptor
LMP	Low molecular weight proteins		
LRT	Low Resolution Typing		

Amino Acid Abbreviations.

Amino Acid	Three Letter	One Letter
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic Acid	Asp	D
Cysteine	Cys	C
Glutamic Acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

Chapter 1

Introduction;
Definition of HLA polymorphism.

1.1 The HLA system.

1.1.1 Introduction

The Major Histocompatibility Complex (MHC) is a central and important component of the immune system. In humans, the MHC or Human Leukocyte Antigen (HLA) region represents a cluster genes on the short arm of chromosome 6 (figure 1.1). Many of these genes have been shown to be involved in the regulation, processing and presentation of antigen in the immune response. The characterisation of the HLA system has revealed a high level of polymorphism amongst the component genes. The determination of this polymorphism has proven important with regard to understanding histocompatibility and the cellular immune response. This chapter discusses the HLA system and the different approaches for its definition, the development and application of one approach, forming the basis of this thesis.

The need for HLA matching donor and recipient for organ transplantation has driven the requirement for identification of specificities in the HLA system. As the capability for testing evolved, firstly through detection of HLA molecules at the cell surface and more recently through typing HLA by DNA analysis, so the degree and extent of polymorphism found in the HLA

system has grown. Ultimately, the complexity of polymorphism within the HLA system has been unravelled by DNA sequencing the genes that translate into the many expressed HLA specificities.

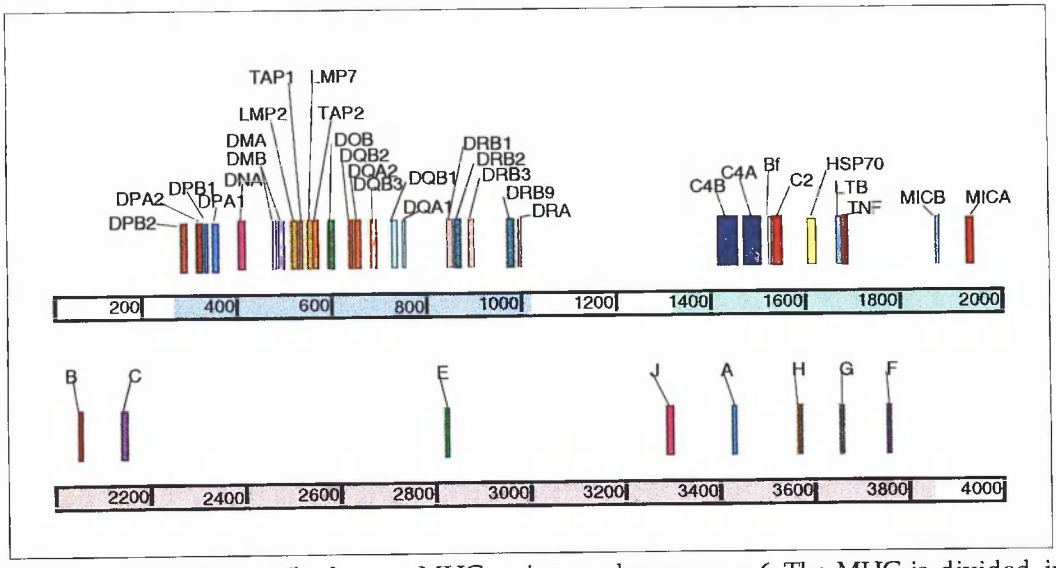


Figure 1.1. A map of the 4Mbp human MHC region on chromosome 6. The MHC is divided into class I (pink), class II (blue) and class III (green) regions. Class II is at the centromeric end, whilst class I is at the telomeric end of the MHC complex. The class I region contains genes encoding amongst other things, the classical HLA-A, -B, -C antigens. The class II region contains genes encoding HLA-DP, DQ, DR antigens. The class III region contains genes encoding some of the components of the complement system, including C2 C4 and Bf.

HLA genes are encoded on the short arm of chromosome 6 and represent the most polymorphic loci known in man (figure 1.1). In 1987, 12 Class I alleles and 9 Class II alleles had been sequenced (Bodmer, J.G. et al 1990). This has quickly grown in number so that in the 1995 Nomenclature report (Bodmer, J.G. et al 1995), some 213 Class I (A,B,C) and 256 Class II (DR,DP,DQ) alleles had been identified. The initial surge of DNA sequencing was biased towards HLA Class II alleles, probably since serological definition of Class II antigens was acknowledged as inadequate for clinical purposes. This allowed DNA based systems to be developed for Class II, which was particularly desirable because of the difficulties in performing Class II typing through serology. The revolution of DNA based technology and its application to identify polymorphism at the nucleotide level provided more information as to the relevance of polymorphism in terms of HLA structure and function. Detection of HLA through DNA analysis achieved an

increased level of resolution, not possible through serological or biochemical means. Characterisation of HLA genes by DNA sequencing and the collection of these into an accessible database (Arnett and Parham, 1995; Marsh, et al 1995) provided the information for the further design of DNA based methods. In turn these methods identified new variants, which when sequenced consolidate and add to the sequence information already held on HLA polymorphism.

Elucidation of the crystal structure of the Class I and Class II molecules (Bjorkman, et al 1987; Gorga, et al 1991; Saper, et al 1991; Brown, et al 1993) provided a physical model within which the function of HLA polymorphism could be better understood. The high level of polymorphism found around the peptide binding groove, determines the nature and conformation of the peptide bound by the HLA molecule (Saper, et al 1991; Kubo, et al 1994; Tanigaki, et al 1994). Therefore, HLA polymorphism has implications for antigen presentation in the T cell mediated immune response (Gotch, et al 1985; Latron, et al 1991; Rotzschke, et al 1992; Utz, et al 1992; Tussey, et al 1994). The significance of these polymorphic differences has become increasingly apparent as more HLA specificities have been identified.

The HLA system is of major importance in the functioning of the immune response. The polymorphic nature of the HLA system and its function has prompted the application of tissue typing techniques in a number of areas, including transplantation, epidemiology and disease association, anthropology and cellular immunology. To understand better the relevance of HLA polymorphism with regard to the immune response, it is first necessary to develop accurate methods for determination of HLA specificities. This thesis charts the development of one such approach in application to the HLA-A locus. It describes the initial development of a basic PCR based approach and the subsequent pursuit of methods for determining an allelic level of resolution at the HLA-A locus. Through this work, the feasibility of high resolution HLA typing is realised.

Detection of HLA polymorphism has utilised a wide array of techniques over the years. The observations made in identifying polymorphic differences by different methods, have in turn lead to a greater

understanding of the nature of the HLA system. This introduction will briefly describe the HLA system and how different approaches have been applied to detection of polymorphism both for Class I and Class II loci.

1.1.2. HLA stucture and polymorphism.

HLA Class I and Class II antigens are highly polymorphic molecules expressed at the cell surface. HLA Class I molecules comprise a polymorphic heavy chain associated with a β2-microglobulin invariant light chain and endogenously processed peptide (Bjorkman, et al 1987; Saper, et al 1991). HLA Class II molecules comprise two covalently linked heavy chains associated with peptide (Brown, et al 1993). The heavy chains for Class I and Class II molecules have transmembrane and cytoplasmic regions. The genes coding for the HLA heavy chains are found on the short arm of chromosome 6 (figure 1.1). The gene coding for the human β2-microglobulin light chain is located on chromosome 15.

HLA Class I antigens are generated from three classical loci, HLA-A, -B, -C. These classical Class I molecules are expressed at the surface of most nucleated cells and act as antigen presenting moleclues for $CD8^+$ T cells. Some non-classical HLA Class I antigens also exist, namely HLA-E, -F, -G which also associate with β2-microglobulin. These non classical HLA molecules have restricted tissue expression. HLA-G is predominantly expressed on the trophoblast (Kovats, et al 1990; Rinke, et al 1990).

Classical HLA Class II antigens comprise HLA-DR, -DQ, -DP. The HLA-DR antigens comprise a polymorphic beta chain in association with a invariant alpha chain. In addition to the DRB1 locus which codes for most of the HLA-DR polymorphism, the DR alpha invariant chain can combine with beta chains coded for at three other DR loci, namely DRB3, DRB4 and DRB5 to give antigens DR52, DR53 and DR51 respectively. HLA-DQ and DP comprise alpha and beta chains, both containing polymorphism. Class II antigens are restricted in their distribution, and are expressed on B cells, macrophages, monocytes, activated T cells and a few additional cell types. HLA Class II molecules present antigen to $CD4^+$ T cells.

The function of HLA molecules became better understood through the elucidation of their structure. This was first achieved through the determination of the crystal structure of the HLA-A2 molecule (Bjorkman, et al 1987). This revealed the HLA-A2 molecule consisted of a peptide binding groove (formed by the alpha-1 and alpha-2 domains) supported by the combination of the alpha-3 domain and the β2-microglobulin light chain (figure 1.2). A similar structure was noted for Class II (Brown, et al 1993), with the peptide binding groove comprising the alpha-1 and beta-1 domains. The main regions of sequence hypervariability are found around the peptide binding groove reflecting their function for HLA restricted antigen presentation. Polymorphism within the peptide binding groove affects the sequence and conformation of peptide bound within the groove (Tussey, et al 1994).

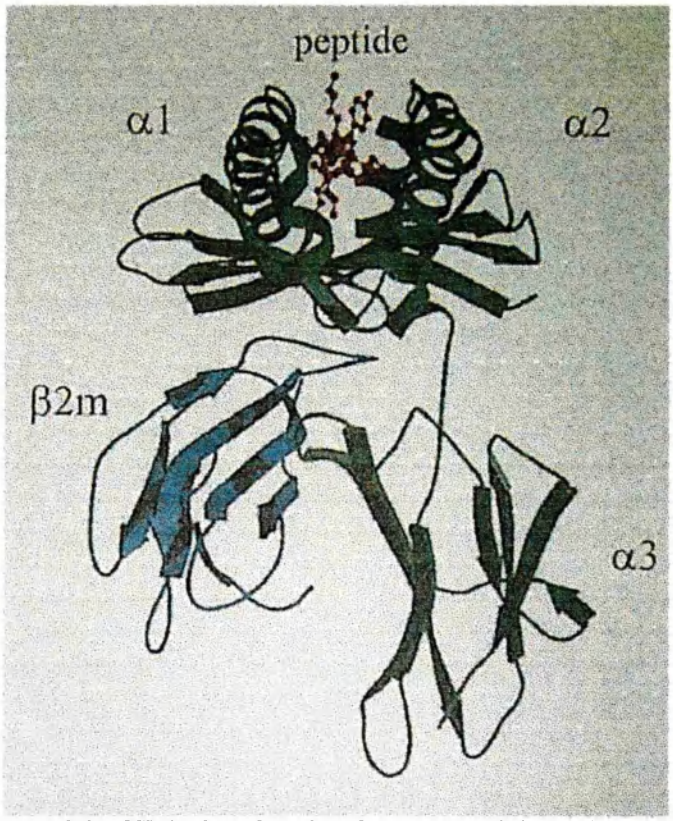


Figure 1.2. A depiction of the HLA class I molecule as viewed from the side looking down the peptide binding groove. The peptide is shown in red. The peptide binding groove is formed by the α-1 and α-2 domains of the class I heavy chain (green). The binding groove is supported by a platform formed by the α-3 domain in association with the β-2microglobulin light chain shown in blue. This model, based on the HLA-B8 molecule was generated in MOLSCRIPT (Kraulis, 1991, modification by R.Esnouf), and rendered with Raster3D (Merritt, et al 1994).

The majority of HLA Class I and Class II sequence polymorphism identified at the nucleotide level are coding differences, translating into amino acid substitutions. Very few sequence polymorphisms are silent and most polymorphic differences are located in or around the peptide binding groove. These observations imply the evolution and maintenace of HLA polymorphism through selection mediated by pathogenic pressure (Lawlor, et al 1990). Additionally, polymorphism has also been observed in the promoter regions (Cereb, et al 1994) and intronic regions of the HLA sequences (Cereb, et al 1995). Both may be useful for DNA based tissue typing.

A large number of sequences defining HLA specificities have now been published (Arnett and Parham, 1995). In comparing the HLA sequences, the polymorphic differences are in the most part not represented as unique sequence motifs. Instead, most alleles gain their uniqueness through a mosaic of shared polymorphic differences, each defined at a particular hypervariable region (see figure 1.10). The nature of this type of polymorphism seems to be the result of gene conversion, recombination or exon shuffling events (Parham, et al 1988; Erlich, et al 1991b; Geraghty, et al 1992; Hildebrand, et al 1992; Kuhner, et al 1992; Zemmour, et al 1992a; Hildebrand, et al 1994). The combination of these polymorphic differences along the length of the gene (in particular exons 2 and 3 for class I and exon 2 for Class II) distinguish each antigen.

The importance of polymorphism within the HLA system bears significance as to the function of the expressed molecule and how it mediates the immune response in terms of antigen presentation. The fact that the majority of polymorphism in HLA is found around the peptide binding groove underlines the requirements for identification of these sequence differences. This can be achieved by two main approaches; detection of the expressed molecule or detection of the genes.

1.1.3. MHC restriction, antigen processing and antigen presentation by HLA class I molecules.

The HLA class I molecule is a central player in the cellular immune response, acting as an alloantigen in the mixed lymphocyte reaction (discussed below 1.2.3) and as a restriction element for recognition of antigen by the cytotoxic T cells (CTL). Evidence for the use of MHC class I as a restrictive element was noted in murine models (Zinkernagel 1974; Zinkernagel and Doherty, 1974a; Zinkernagel and Doherty, 1974b). This work demonstrated that virally infected target cells were lysed by specific CTL which shared at least one mouse class I MHC (H-2) antigenic specificity. CTL which did not share a H-2 specificity with the targets, failed to lyse the cells, even though they were infected with the same virus to which the CTL were raised.

It was thought that these restriction phenomena could be explained by one of two theories. Firstly, that there was a physiological interaction, in which T cells could recognise antigen bound to the cell surface, but for lysis to occur, an additional self to self signal was required (Miller, et al 1976). The second theory conjectured that lysis required a specific T cell receptor which could identify altered self structures coded for within the MHC genes.

The evidence, although not conclusive, seemed to favour the latter altered self hypothesis. Experiments using bone marrow chimeras in mice showed the presence of two populations of CTL, one specific for syngeneic virally infected cells, the other exhibiting cytotoxicity against the tolerated allogeneic virally infected cells. The specificity of the CTL was shown to be restricted through the H-2 type to which they were raised, and failed to crossreact with the other MHC components in the chimeric model. Both groups of CTL failed to lyse infected cells of an unrelated H-2 type (Zinkernagel, 1976). These experiments showed that the recognition of MHC is in one direction, with the T cell receptor specifically recognising virally altered MHC. This was supported by observations made in blocking experiments where, antisera bound to H-2 specificities on the targets and not the CTL abrogated specific lysis (Germain, et al 1975; Schrader and Edelman, 1976). This further affirmed the notion that the T cell receptor detected MHC which had become altered by the viral infection of the target cell.

Elucidation of the process MHC restriction of cell lysis proceeded further to determine the nature of altered self.

Further investigation of CTL recognition of influenza nucleoproteins transfected into mouse fibroblast cells, provided additional insight into the nature of the MHC restricted immune interaction (Townsend, et al 1984). The recognition of nucleoprotein, a non membrane component of the virus which accumulates in the nucleus of the infected cell, differed with the assumed belief that CTL recognised antigens in their native form on the cell surface, with T helper cells recognising denatured fragments on professional antigen presenting cells (Unanue, et al 1984). This begged the question of how the nucleoprotein reached the cell surface. An explanation of the results (Townsend, et al 1984), suggested that other cells including the transfected fibroblasts, could also process and present viral antigens, provided they were generated within the cell. This observation explained previous results where antibodies specific for native viral proteins failed to diminish CTL recognition (Zinkernagel and Rosenthal, 1981), suggesting some type of processing of antigen prior to its expression at the cell surface.

To further investigate the CTL recognition of influenza nucleoprotein, deletion mutants of the nucleoprotein were transfected into mouse fibroblast cells (Townsend, et al 1985). The results showed that fragments of the nucleoprotein were transported to the cell surface and recognised by CTL. The data also revealed that different CTL clones were able to show specificity for different regions of the nucleoprotein. The experiments also demonstrated that both ends of the nucleoprotein could be independently transported to the cell surface. Antibodies raised against the native influenza nucleoprotein again failed to abolish CTL recognition. Additionally, antibody against the trunctated nucleoprotein could only detect very small quantities at the cell surface, which was contradictory to the CTL data. These results suggested that viral antigen had been denatured prior to surface expression, perhaps as part of the normal cytosolic processes of protein degradation.

Experiments then followed to determine whether CTL did indeed recognise short fragments of the antigenic protein in a fashion similar to recognition of processed antigen by class II restricted T helper cells. This involved the

generation of short synthetic peptides defining epitopes within the influenza nucleoprotein (Townsend, et al 1986). The results showed clearly that short peptides representing epitopes of the influenza nucleoprotein could be recognised by MHC class I restricted CTL. This further enhanced the view that viral proteins required degradation within the cytosol prior to presentation at the cell surface in association with the class I molecule.

Hence it was shown that the HLA class I molecules present endogenously derived peptide at the cell surface for surveillance by the immune system. CTL will normally mount immune responses against cells MHC presenting foreign peptide derived from viral infections or malignant transformations, while cells expressing peptides derived from the normal self proteins will be ignored. To allow presentation of self or foreign peptide requires the degradation and processing of the respective proteins within the cell. The mechanisms allowing presentation of peptide in combination with HLA molecules at the cell surface, have been increasingly studied over recent years. The intracellular mechanisms of antigen processing and assembly of class I molecules have recently been reviewed (Brodsky, et al 1996; Williams, et al 1996) and are outlined below.

The stability of the class I molecule requires the formation of a trimeric structure comprising class I heavy chain, β2-microglobulin and peptide (Townsend, et al 1989). The HLA molecule binds peptide in the region of 8-10 amino acids in length (Rotzschke, et al 1990). The HLA molecule also prefers to bind peptides with certain amino acid residues at particular points within their sequence (Falk, et al 1991). The binding of a preferred peptide motif is more extensively described in chapters 7 and 8.

To achieve the presentation of antigen at the cell surface therefore requires the degradation of endogenous protein, the combination of the class I heavy chain with β2-microglobulin, and subsequent loading of peptide. The assembly of class I molecules occurs in the endoplasmic reticulum (ER), yet the degradation of protein occurs essentially in the cytosol (Engelhard, 1994). This is supported by the discovery of an ATP dependent transporter (TAP), which actively takes peptide from the cytosol into the ER (Townsend and Trowsdale, 1993).

The degradation of peptide in the cytosol has been associated with two proteasomes (20S and 26S) (Peters, 1994). These two complexes are associated with the major proteolytic activity within the cytosol. The 20S proteosome forms the core of the 26S proteosome in a barrel type structure. These proteosomes degrade either ubiquitin-conjugated proteins or denatured proteins. The link with antigen processing and the proteosome was made when two genes coding for two proteosome subunits were identified within the MHC complex (Glynne, et al 1991; Kelly, et al 1991). The incorporation of these two sub-units into the proteosome complex, alters its cleavage characteristics. Expression of these genes known as LMP2 and LMP7, together with others implicated in antigen processing, is inducible by γ-inteferon. Good evidence for the contribution of LMP2 and LMP7 to antigen processing is that cells which lack the genes express reduced levels of surface class I molecules. This can be reversed by the addition of exogenous peptide (Fehling, et al 1994). It is also possible that proteosomes can to some extent determine the range of peptide motifs available for presentation. If the proteosome has a preferred cleavage site at the centre of a possible motif, then this potential peptide would not be available for presentation (Niedermann, et al 1995).

The proteosomes degrade protein to produce peptides in the range of 3-15 amino acids in length, consistent with that required for loading onto and presentation by the class I molecule. They are also of the correct size for transport into the ER by the TAP molecule, which exhibits a preference for peptides in the region of 8-15 amino acids in length (Howard, 1995). The transport of peptide into the ER was clarified through the discovery of two genes within the MHC complex, near to the LMP2 and LMP7. These two genes, (TAP1 and TAP2) encode the two subunits of an ATP dependent transporter molecule (Trowsdale, et al 1990; Kleijmeer, et al 1992). Cells lacking TAP genes have low expression of HLA class I molecules at the cell surface and an accumulation of unstable heavy chain/β2-microglobulin complexes within the cell (Howard, 1995). This evidence seems to underline the importance of the TAP molecule in transportation of peptide from the cytosol into the ER facilitaing stabalisation of the class I molecule.

So far there has only been a limited amount of polymorphism observed in the human TAP genes. There is however significant polymorphism

observed amongst the rat TAP genes, which affect the specificity of the TAP molecule in its transportation of peptide. In the mouse, the TAP molecule shows a preference for peptides with a hydrophobic C-termini, which is reflected in the peptide motifs bound by the murine class I molecules. The human TAP does not seem to have this level of preference, although some peptides are poorly transported (Powis, et al 1992).

Once peptide enters the ER, it may bind to the class I heavy chain/β2-microglobulin complex. If the peptide is longer than that preferred, additional trimming can occur within the ER. It is unclear whether this trimming occurs before or after loading of peptide onto the class I molecule. There is also evidence of binding of peptide to a chaperone molecule (gp96) within the ER (Srivastava, et al 1994). It appears that this chaperone can protect against rapid degradation of the pepide before it is loaded into the empty class I molecule. Peptide can also be recycled from the ER into the cytosol by an as yet undefined ATP dependent pathway (Brodsky, et al 1996).

The loading of peptide onto the class I molecule requires the association of the heavy chain with β2-microglobulin. It has also been shown that the heavy chain/β2-microglobulin complex associates with the TAP molecule (Ortmann, et al 1994; Suh, et al 1994), with the interaction occuring through the TAP1 subunit. This association may have evolved to accommodate more efficient loading of peptide onto the empty class I molecule.

Before peptide can be loaded onto the empty class I molecule, the heavy chain needs to be correctly folded and assembled with β2-microglobulin. The assembly within the ER has been associated with a number of chaperone molecules, the most characterised being calnexin (Helenius, 1994; Williams, 1995). This chaperone has been shown to assist the folding of the class I heavy chain and incorporation of the β2-microglobulin light chain. The calnexin-HLA complex then associates with TAP. Disassociation from TAP occurs once peptide is bound (Williams, et al 1996). The point at which calnexin disassociates from the HLA molecule is unclear. It has also been shown that calnexin deficient cells can still express HLA on their surface, which indicates the presence of other chaperones which facilitate the assembly of class I molecules (Scott and Dawson, 1995).

Following disassociation from TAP and calnexin, the class I molecule is released from the ER and is transported to the cell surface via the golgi apparatus (Brodsky et al. 1996; Williams et al. 1996). Further work is required to elucidate the complete picture of antigen processing and presentation. However enough is now known to consider the interactions and mechanisms as complex. A simplified depiction of the events mentioned above is given in figure 1.3.

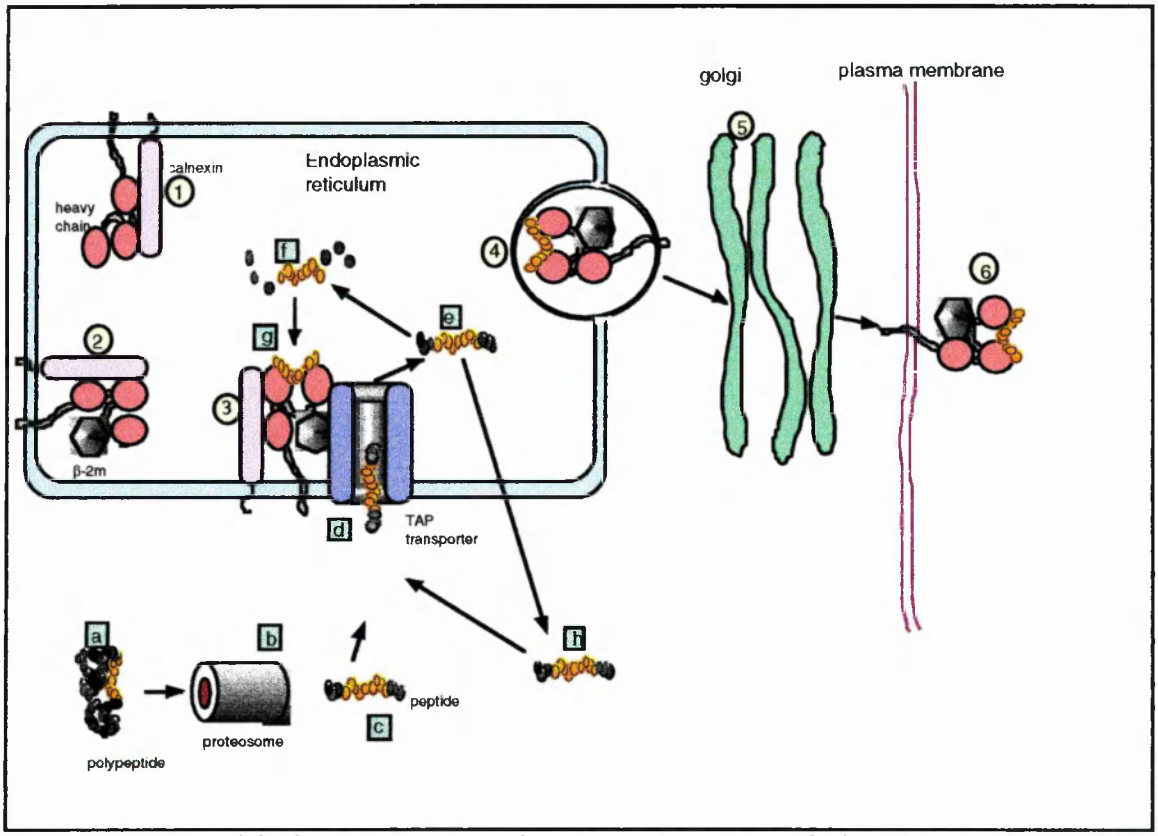


Figure 1.3. A simplified representation of antigen processing and class I assembly. Antigen processing. Protein (a) containing the relevant peptide is cut by the proteosome (b) in the cytosol. The resulting peptides (c) are transported into the endoplasmic reticulum (ER) by the TAP transporters (d). In the ER, the peptide (e) can be trimmed further (f) prior to loading onto the class I molecule (g) or it can be recycled back to the cytosol (h). It is unclear whether trimming of the peptide occurs before or after it is bound to the class I molecule. Class I assembly. Nascent heavy chain is bound by calnexin (1). This allows binding of β2-microglobulin (2). When this complex is bound to the TAP transporter, peptide can be loaded onto the class I molecule (3). Binding of peptide triggers disassociation from TAP and at some point calnexin (4). The class I molecule is then transported through the golgi (5) and expressed at the cell surface (6).

1.1.4 Nomenclature.

There have been a number of different techniques used to define HLA polyphorphism, including serological, biochemical, T cell and molecular biological methods. These methods gauge the differences present in HLA genes or their products from different perspectives. Different methods each had their own nomenclature in identifying the specificities they were determining.

Initially, HLA specificities were defined by serology and named through a combination of a number in conjunction with the gene locus. (e.g. A10; DR2). As an increasing array of sera were used and cells tested, so it became possible to split many of the initial specificities into a number of serologically defined antigens. Hence A10 could be split into A25, A26, A43, A34 and A66, and DR2 could be split into DR15 and DR16. Additionally, the use of Isoelectric focusing (IEF) could further split these serologically defined specificities into subtypes, and they were named by their serological specificity followed by a number (e.g. A2 could be subtyped by IEF into subtypes A2.1, A2.2, A2.3, A2.4, A2.5). A list of HLA-A alleles, their broad serological specificities and splits, and related IEF specificities, are shown in Table 1.1. To confuse matters further, another nomenclature was required to define specificities determined through T cell recognition of allogeneic HLA, with particular reference to Class II specificities (e.g., serologically defined DR4 could be subdivided into Dw4, Dw10, Dw13, Dw14 and Dw15 by specific T cell clones).

This complex situation was resolved through the introduction of a nomenclature system based on the allelic definition of HLA specificities through the DNA sequence of their genes. This system required that alleles were defined by their gene locus in combination with a four digit number; the first two digits relating to the specificity (mainly on the basis of serology), and the last two digits relate to the subtype number. Hence, A*0205 denotes the fifth HLA-A2 allele. A fifth digit has been introduced to allow distinction between alleles containing silent polymorphisms. This system of nomenclature, sanctioned by the World Health Organisation, provides an

Allele	Serological Specificity	IEF position	Local Name
A*0101	A1	A1	-
A*0102	A1	-	-
A*0201	A2	A2.2	A2.1
A*0202	A2	A2.1	A2.2F
A*0203	A203	A2.3	A2.3
A*0204	A2	A2.4	A2.4
A*0205	A2	A2.1	A2.2Y
A*0206	A2	A2.2	A2.4a
A*0207	A2	A2.2	A2.4b
A*0208	A2	A2.2	A2.4c
A*0209	A2	A2.3	A2-OZB
A*0210	A210	A2.2	A2-Lee
A*0211	A2	A2.5	A2.5
A*0212	A2	A2.2	-
A*0213	A2		A2-SLU
A*0214	A2	-	A2-1S
A*0215N	-	-	A2-Null
A*0216	A2	-	A2-TUB
A*0217	A2	-	A*New
A*0301	A3	A3.2	A3.1
A*0302	A3	-	A3.2
A*1101	A11	A11.2	A11E
A*1102	A11	-	A11K
A*2301	A23(A9)	A23	-
A*2402	A24(A9)	A24.1	-
A*2403	A24(A9	A24.1	A9.3
A*2404	A24(A9)	-	A24AK
A*2405	A24(A9)	-	A24New
A*2406	A24(A9)	-	A*24YM
A*2501	A25(A10)	A25	-
A*2601	A26(A10)	A26.1	A26.1 A26.3
A*2602	A26(A10)	-	A26.2 A26.1
A*2603	A26(A10)	-	A26.4
A*2604	A26(A10)	-	A10SA
A*2901	A29(A19)	A29.1	-
A*2902	A29(A19)	A29.2	A29.2
A*3001	A30(A19)	A30.3	A30.3
A*3002	A30(A19)	A30.1	A30.2
A*3003	A30(A19)	-	A30JS
A*3004	A30(A19)	A30.2	A30W7
A*31011	A31(A19)	A31	A31
A*31012	A31(A19)	A31	-
A*3201	A32(A19)	-	-
A*3301	A33(A19)	A33.2	Aw33.1
A*3302	A33(A19)	-	Aw33.2
A*3303	A33(A19)	-	A33NC
A*3401	A34(A10)	A34.1	-
A*3402	A34(A10)	A34.2	-
A*3601	A36	-	-
A*4301	A43(A10)	-	-
A*6601	A66(A10)	A26.2	A66.1
A*6602	A66(A10)	-	A66.2
A*68011	A68(A28)	A28.2	Aw68.1
A*68012	A68(A28)	A28.2	Aw68.1
A*6802	A68(A28)	A28.1	Aw68.1
A*6901	A69(A28)	A28.1	-
A*7401	A74(A19)	-	-
A*8001	-	-	AX'BG'

Table 1.1. List of HLA-A locus alleles adapted from the 1995 WHO Nomenclature report (Bodmer, 1995). This table also lists the serological, IEF and local names given to the HLA-A locus specificities. Serological types in brackets denote the broad specificities from which the given specificity was split.

unambiguous assignment of HLA specificities within the parameters of known HLA alleles.

1.2 Detection of polymorphism in HLA molecules.

1.2.1 Serological determination.

The HLA system was discovered through observations of agglutination of leukocytes with antisera taken from unrelated transfused individuals (Dausset, 1954). The original specificity found (Dausset, 1954) (Mac) was later to be known as HLA-A2. Similarly, specificities relating to Bw4, Bw6 (van Rood and van Leeuwen, 1963) were identified, followed by A1 and A3 (Payne, et al. 1964) and this marked the unfolding of the polymorphic nature of the HLA system.

The ability to detect these antigens advanced through the use of the HLA specific antisera for complement mediated cytolysis of HLA expressing leukocytes (Terasaki, et al 1964; Bodmer, W.F. et al 1967; Kissmeyer, et al 1969). This microlymphocytotoxicity assay became the standard approach to histocompatibility testing from the 1960s through four decades. The use of polyclonal antisera from transfused individuals, or more commonly from maternal blood (antisera present against paternal HLA determinants), provided the means for identification for the increasing polymorphism found within the HLA system.

As shown in figure 1.4, serological HLA typing (also known as the microlymphocytotoxicity assay), detects expressed HLA antigens through the specificity of antisera and their ability to initiate complement mediated cellular lysis. Cellular lysis, as detected through the uptake of an appropriate dye, confers the specificity of the typing antisera on the sample tested. A panel of well characterised antisera, is used to detect all known HLA specificities. The panel of antisera is incubated with seperated peripheral blood mononuclear cells in Terasaki plates. Complement is added to the cells and antisera, and the mixture incubated for a further period. Dye is added and a positive reaction with an antisera is detected through the diffusion of this dye into the lysed cell. If the antisera does not have

specificity for the HLA expressed on the cell, then lysis does not occur and dye is not seen within the cell. Because of their polyclonal and crossreactive nature, many antisera are necessary in a typing panel and tissue type determined through reaction patterns.

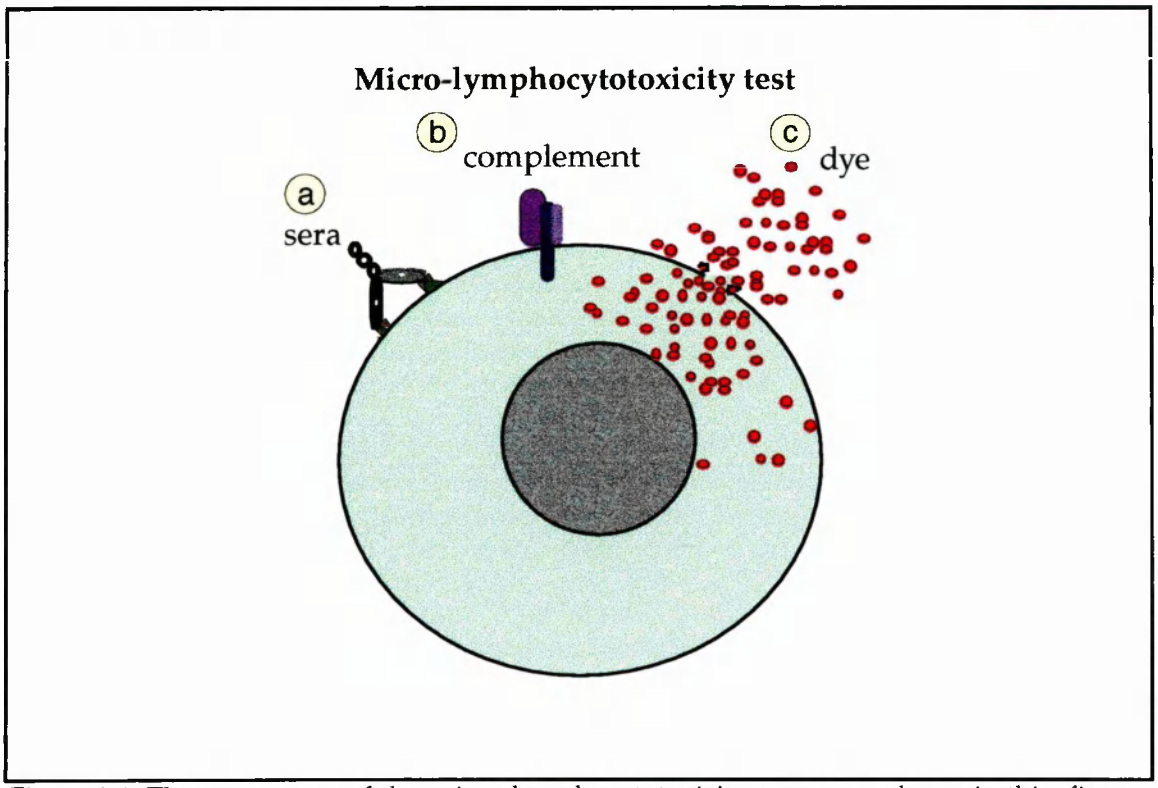


Figure 1.4. The components of the micro-lymphocytotoxicity assay are shown in this figure. Viable peripheral blood mononuclear cells (Class I) or separated B cells (Class II) are mixed with antisera of known HLA specificity. If antisera is bound to a cell (a), this mediates the activation of complement, which is added to the assay. Complement in the presence of the appropriate immunoglobulin class of antisera will puncture the cell (b). Dye can then enter the dead cell denoting a positive result (c), and this can be visually detected through an appropriate microscope. The absence of dye within the cell denotes a negative result and indicates the specificity of the antisera is different to that expressed by the cells tested.

Identification of HLA Class I antigens improved as more antisera were tested against larger panels of antigens. Such serological characterisation uncovered more and more polymorphism as initially characterised broad specificities were increasingly split. For example, at the HLA-A locus, HLA-A9 was spilt into A23 and A24, and HLA-A19 was split into A29, A30, A31, A32, A33, A74. Similarly HLA-B and HLA-C specificities were split through their reactivity with panels of antisera.

Some of this characterisation was made easier through the use of homozygous typing cell lines. These cell lines which have homozygous haplotypes were obtained from individuals resulting from consanguinous marriages. The use of panels of such cell lines allows the specificity of multispecific antisera to be more readily distinguished.

Serological characterisation of antigens and antisera derived from different ethnic groups observed both new specificities and different antigen frequencies from populations previously investigated. Many such studies were organised through International Histocompatibility Workshops. They provided a large number antisera which were tested against cell line panels collected from different populations. This exercise provided an extremely useful resource, since antisera and cell lines made available through the workshop were subsequently highly characterised. In addition, recent workshops have organised 'antigen societies' which intensively characterised groups of crossreactive HLA antigens. The antigen societies searched for novel specificities and identified any linkage disequilibrium prevalent with the antigens of interest. They also identified key antisera for identification of new specificities, together with their ethnic distribution.

Serology, particularly for Class I, remains the routine mainstay method for many histocompatibilty testing laboratories. Developments over the years, such as the use of two colour fluoresence (Bodmer, W.F. et al 1979), to more readily detect cell lysis, and varying degrees of automation, have allowed improvements in the basic method. HLA typing by serology does require a certain skill. The serologist needs to be aware of the specificity of the antisera in his typing panel, and be able to identify rare HLA specificities, which may not be immediately apparent from the reactivity obtained with a typing panel. Serology is handicapped by its reliance on viable cells. The lower the viability, the higher the background which makes detection of complement mediated cellular lysis more difficult.

Serology for Class II specificities (HLA-DR, -DP, -DQ) was developed at a later date (van Leeuwen, et al 1973; Ceppellini, et al 1974; Jones, et al 1975; Bodmer, J. et al 1976; Bodmer, J. 1978; van Leeuwen, et al 1982). HLA Class II antigens were discovered through observations made in the mixed lymphocyte test (Amos, et al 1971; Yunis, et al 1971; Mempel, et al 1972). This

pointed to another component in the HLA system in addition to HLA-A, -B, -C. Identification of Class II specificities were then achieved through serological identification, using sera which had been treated with platelets (expressing Class I but not Class II) to remove the presence and so interference of the Class I component in the polyclonal antisera. The use of Class I negative cell lines such as Daudi (de Preval, et al 1983; Rosa, et al 1983; Browning, et al 1995) provided additional information as to the organisation of Class II and the understanding that more than one locus was involved (Bodmer, W. et al 1984). Further analysis through amino acid sequencing identified Class II molecules as heterodimers comprised of an α and β chain. The genes for these chains are all found in the HLA cluster of genes on chromosome 6. In DR specificities, the polymorphic DRB1 (and DRB3, DRB4 and DRB5) provide the β chain, and this complexes with the invariant α chain (coded by the DRA1 gene). Sequence polymorphism in DQ and DP can be seen in both α and β chains.

The molecular analysis of HLA Class II provided the means for understanding serological observations. The interpretation of the specificity of antisera could be made in the context of antigenic sequence polymorphism found within the HLA molecule. Hence an antisera which identifies a group of specificities may be detecting a polymorphism common to their respective sequences. As the means to generate monoclonal antibodies to HLA specificties became possible, a clearer picture emerged as to the epitopes the antibodies were detecting (Marsh, et al 1989).

The use of monoclonal antibodies in the field of histocompatibility testing promised a renewable source of monospecific reagents which would circumvent the problem of crossreactivity present in polyclonal antisera (Barnstable, et al 1978; Brodsky, et al 1979; Bodmer, J. et al 1989). However the cross-reactivity of antisera could partially be explained in another way, particularly with the increasing amount of sequence data available. Polyclonal antisera contain multiple specificities directed against different antigenic determinants on one or more of the immunising HLA molecules. However, as discussed previously, most of these antigenic epitopes or 'motifs' are not peculiar to a particular HLA specificity, but can be shared by several. For instance, a monoclonal antibody has been described which has specificity for A2 (but not A*0210) and A69 (Parham, et al 1981). The

specificity of this monoclonal antibody can be interpreted by looking at the sequences of A2 and A69 alleles, with the epitope mapped to a tryptophan at amino acid position 107. This is at a position which is accessible for antibody detection. Therefore monoclonal antibodies which identify a single serologically defined HLA specificity are rare, and typing still relies on reaction patterns. But the level of crossreactivity is decreased, reducing the size of the typing panel and making interpretation easier. Typing plates using monoclonal antibodies have been compiled for Class I and Class II typing and are commercially available.

Serological identification of Class II has proved more difficult than Class I. Firstly, only the B cells and monocytes, which represent a small proportion of peripheral blood mononuclear cells (PBMC), express HLA Class II. Therefore, to perform Class II typing, separation of B cells from the PBMC population requires a sufficient volume of blood. Alternatively, B cells can be immortalised through EBV transformation, but this does not represent a practical approach to routine Class II typing due to the length of time this procedure takes. Determining HLA Class II specificities by serology therefore has an added layer of complexity compared to typing Class I.

The difficulty in typing for HLA Class II through serological means provided the incentive for adopting new approaches to histocompatibility testing, namely through DNA analysis. Serological typing for Class I was however a more routine affair and there was no great pressure to find alternative means for typing. Also the lack of Class I alleles sequence data, prohibited the development of DNA based typing for these specificities.

Serological typing can only assess those epitopes accessible to antibody. Polymorphism not detected by serology was demonstrated by IEF analysis (Biddison, et al 1982) and through allo-reactive T cell discrimination (Gotch, et al 1985). The latter showed that the resolution offered by serology was potentially inadequate as it failed to identify Class I specificities to the same level as T cells, with obvious implications for transplantation matching. To this end, additional methods of characterisation had to be considered.

1.2.2 Biochemical detection of HLA expression.

One dimensional isoelectric focusing (IEF) (Biddison, et al 1982; Yang, 1989b), detects amino acid substitutions contributing to changes in charge of the HLA molecule. The radio-labelled (usually ^{35}S methionine) class I molecule is immunoprecipitated with an appropriate antibody. Electrophoresis is carried out on the immunoprecipitate through a pH gradient. The immunoprecipitate will focus at a point appropriate to its charge, forming a band on an autoradiograph. An example of an IEF gel is given in chapter 6 (figure 6.3).

The use of IEF for tissue typing has demonstrated a level of polymorphism beyond that observed in serological typing. Reports from International Histocompatibility workshops (Yang, 1989a) have listed IEF variants, identified their relative focal point and described any population and frequency data known. For example, five A2 IEF variants, four A24 variants, and four A30 variants have been identified. Increasing the resolution in the IEF can also show small differences in focal points, as described in characterisation of a new A2 IEF variant (Guttridge, et al 1992) later confirmed through DNA sequencing (Barber, et al 1994).

Unfortunately, IEF may not be used independently. It requires initial information from serology to allow interpretation of the focused bands. This is due to many of the bands focusing in close proximity to each other. So although IEF typing fails to provide a clear and easy to interpret method for tissue typing, it does provide an extension to serology, offering a further level of resolution. Variants not detected through serology can be clearly distinguished through IEF analysis. Identification of these variants allows their further characterisation through DNA sequencing, so that the relevance and nature of polymorphism can be assessed.

1.2.3 Typing through the specificity of T cell recognition.

Concurrent to observations being made using antisera with HLA specificity, it was also noted that lymphocytes from two unrelated sources when mixed and cultured, would transform and proliferate (Hirschorn, et al 1963). This

would not occur when the lymphocytes were obtained from HLA identical sib pairs (Bach, et al 1967). The proliferation noted in this mixed lymphocyte reaction (MLR) could be quantified through incorporation of radioactive ^{3}H-thymidine into the proliferating cells. As illustrated in figure 1.5, the MLR was made easier to interpret through inactivation (using irradiation or Mytomycin-c treatment) of one of the stimulator populations (Bach, et al., 1966). Inactivated stimulator cells that had been well characterised for HLA specificities, could be used as typing cells. The use of homozygous typing cells as stimulators (Mempel, et al., 1973b), delivered a more concise result since only one HLA haplotype need be considered.

The MLR provided an initial technique for determination of HLA Class II specificities. Specificities identified in the MLR were denoted by the 'Dw' nomenclature (Bodmer, W.F. et al 1984) which later showed relation to the serologically defined 'DR' specificities (Jaraquemada, et al 1986). The MLR also provided a method for matching donor with recipient when considering transplantation. Studies have shown that serologically matched unrelated individuals can produce a positive MLR (Mempel, et al 1973a). This indicates either that serology is insufficiently sensitive to perfectly match Class II specificities, or the presence of mismatches amongst other 'minor' histocompatibility antigens, not considered in the serological typing. The MLR was therefore capable of detecting functional polymorphism not seen by serology. This polymorphism identified in the MLR has been shown to be significant in terms of the survival of the graft in the transplant (Morishima, et al 1995). However, negative MLR may not indicate perfect matching, and these undetected mismatches may still contribute to the rejection of the graft (Tiercy, et al 1991; Eiermann, et al 1992).

It was noted that the proliferating T cell clones produced in an MLR were capable of generating a substantive and rapid secondary response when repeatedly restimulated with the HLA antigenic determinant against which they were raised (Sheehy, et al 1975). This observation was adapted into the primed lymphocyte test (PLT). Lymphocytes were cultured for 1-2 weeks in an MLR using HLA typed inactivated stimulator cells. The resulting proliferative lymphocytes could be cloned through limiting dilution and used to type test cells. Proliferation of the PLT clones with the test cells

would indicate the presence of the same HLA specificity as contained in the initial stimulating cell. PLT typing was useful for identifying and characterizing HLA-DP specificities (Morling, et al 1980).

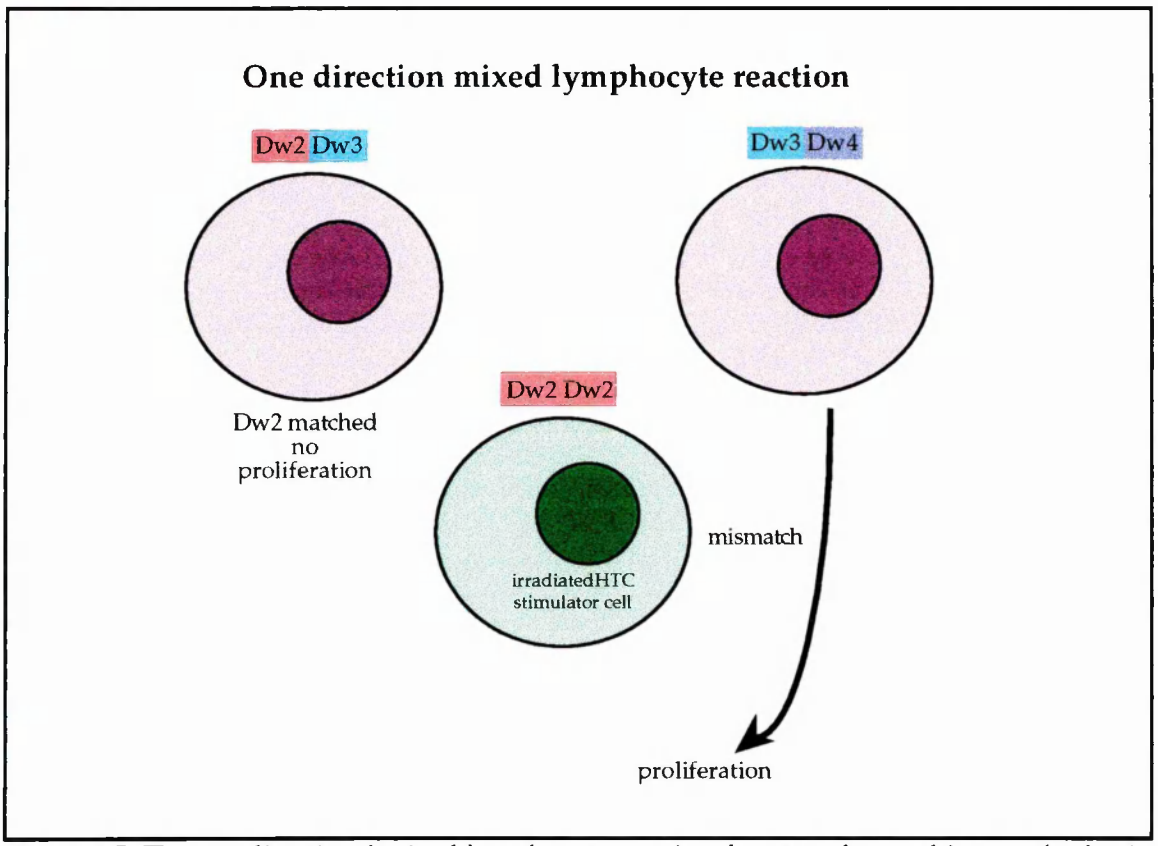


Figure 1.5. The one directional mixed lymphocyte reaction detects polymorphism on the basis that mixing HLA mismatched cells will cause stimulation and proliferation of the cells, the latter detected through incorporation of ^{3}H-thymidine into the mitotic population. The assay can be described as one way, if one of the stimulator cells (of known HLA type), are prevented from proliferation through irradiation or mitomycin-C treatment. Use of HLA characterised homozygous typing cells (HTC) as the stimulator population further simplifies interpretation. In the examples given, the HTC fails to stimulate the Dw2 Dw3 cell through being matched at Dw2. However there is no match with the Dw3 Dw4 cell, allowing its stimulation and proliferation.

CTL clones can specifically recognise and kill a target cell expressing allo-antigen in the context of self-HLA (MHC restriction) (Zinkernagel and Doherty, 1974ab). Such CTL can be specific for viral or tumour antigens. In addition, allo-reactive CTL can also be generated in the MLR, which exhibit specificity for peptide bound to allo-HLA. These cells are produced through activation and expansion of a precursor CTL population found within the unstimulated lymphocyte population. The activated CTL can recognise allo-

MHC directly, or indirectly whereby allo-MHC is processed and presented as peptide by autologous antigen presenting cells (reviewed in Hilton, et al 1996).

Both the direct and indirect types of allo-response can be sensitive to HLA polymorphism, although this is slightly more marked for the self HLA restricted CTL. The generation of such CTL clones provide another 'reagent' for demonstrating and characterizing HLA polymorphism. Again, CTL assays can demonstrate polymorphism not detected through standard serological approaches (McMichael, et al 1988; Castano, et al 1991; Rotzschke, et al 1992; Tussey, et al 1994). More recently, assays to estimate the precursor frequencies of cytotoxic T lymphocytes (CTLp) (Zhang, et al 1989) or helper T lymphocytes (HTLp), have been used to assess the potential response by the recipient to donor allo antigens. Determination of the precursor frequency provides another index by which polymorphism can be detected. Dilutions of recipient cells can be cultured with donor cells to assess the precursor frequency. The precursor frequencies for the allo-reactive T cells is estimated from the presence or absence of a proliferative, IL-2 or cytotoxic response within the context of the recipient cell concentration. A high precursor frequency indicates a mismatch in HLA specificity between the donor and recipient cells. Precursor frequency studies confirm that T cells can discriminate between closely related cross-reactive serological specificities. It also provides an indicator of mismatches between donor and recipient which predict rejection of a transplant graft and graft versus host disease (GVHD) in serologically matched unrelated-donor bone marrow transplants (Kaminski, et al 1991).

CTLp can also be performed using HLA typed target or stimulator cells. These can be cultured with limiting dilutions of responder cells, and the precursor frequency estimated. A low precursor frequency in the responder cells indicates the presence of the same HLA types as in the typed stimulator cells, a high precursor frequency indicates HLA differences. Knowing the HLA type of the stimulator cell can be used to ascertain HLA similarities or differences in the responder cells through the specificity of the T cell.

The fine specificity of allo-reactive T cell clones demonstrates the functional significance of a level of polymorphism not readily discriminated by

serological typing. Failure to identify functionally significant mismatches can cause misinterpretation of cellular assays and increase the chance of rejection in transplantation. T cell clones can however cross-react and they may not make ideal typing reagents for this reason (van der Poel, et al 1986). However, the generation of proliferative and cytotoxic clones from apparently HLA matched stimulator cells, does provide an indicator of the inadequacy of a tissue typing method in not identifying a potentially significant polymorphic difference. These observations highlighted the requirement for attainment of a typing resolution similar to the discrimination of a T cell clone.

1.3 Molecular Biology; determination of polymorphism at the nucleotide level.

1.3.1 Introduction.

Analysis of HLA specificities from DNA provided a different approach to defining their polymorphic differences. Rather than looking at differences in the expressed molecule, polymorphism is characterised at the genetic level. DNA based methods hold particular advantages over serological means of identification. There is no need to rely on viable cells. Reagents are renewable and there is increased potential for standardisation. Recent International Histocompatibility Workshops have been instrumental in encouraging development and implementation of new typing techniques and providing a forum for their standardisation.

The feasibility for development of DNA typing techniques was facilitated by the increasing numbers of HLA alleles which had been sequenced. In general terms, the development of DNA typing Class I alleles has encountered greater problems than those experienced for Class II. One reason for this can be explained in comparing the nature of polymorphism between Class I and Class II. Polymorphism within Class II is contained mainly within defined hypervariable regions in exon 2, making differentiation between alleles readily achievable through hybridisation with the relevant probe. This situation is more complex in Class I alleles. Hypervariable regions are found to different degrees in both exon 2 which

encodes the alpha-1 domain and exon 3 which encodes the alpha-2 domain. (figures 1.6 and 1.10). These two domains form the peptide binding groove of the Class I molecule. Most Class I alleles cannot be differentiated from each other by a particular motif at a particular location in the sequence, but rather the polymorphic differences along the whole length of sequence has to be considered, particularly in exons 2 and 3. This is because, at any particular point of polymorphism, the sequence motif may be shared with not only other alleles within the same locus, but also alleles of other loci and pseudogenes. To identify a Class I specificity by PCR-SSOP therefore generally requires the use of multiple probes, increasing the complexity of hybridisation patterns. Further to this, DNA typing Class I lagged behind Class II, because of the lack of sequenced alleles and the perceived convenience in performing Class I serology.

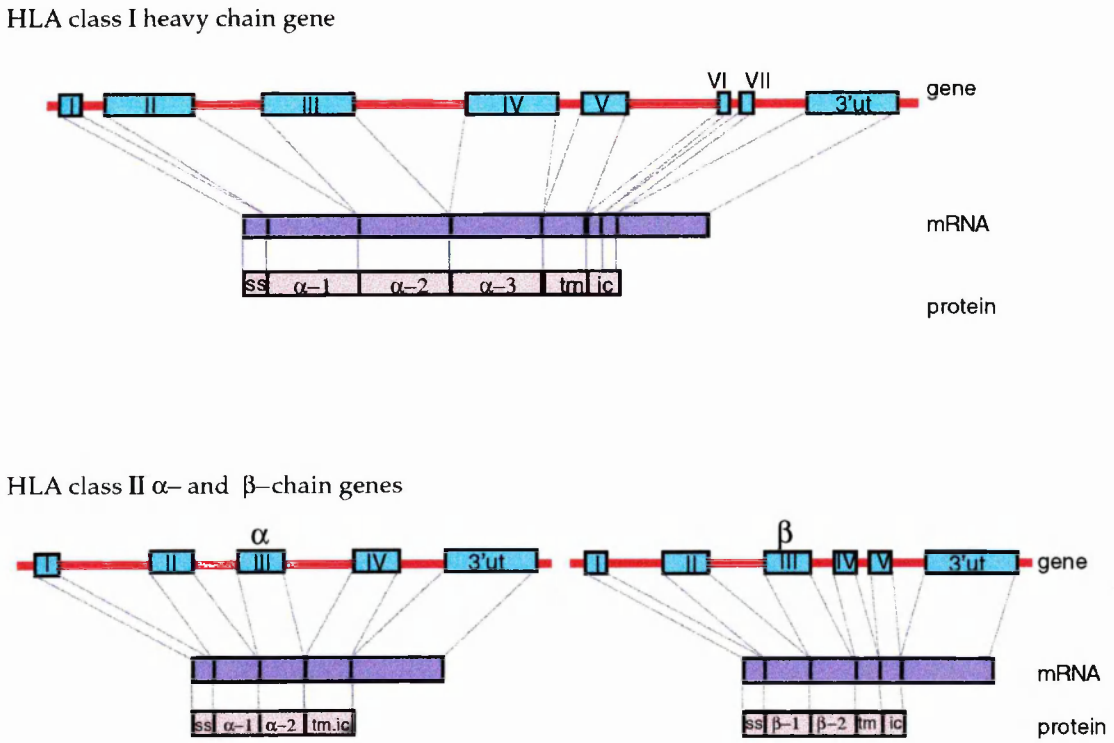


Figure 1.6 The genomic organisation of HLA class I and class II genes and the polypeptide domains of the HLA molecules they encode. (ss = signal sequence, tm = transmembrane domain, ic = intracellular domain, 3'ut = 3' untranslated region).

1.3.2 Restriction fragment length polymorphism (RFLP) analysis.

Initially, molecular biological approaches did not have sufficient information to design highly definitive DNA typing approaches. Because only a few HLA specificities had been sequenced, analysis of DNA was an extension of other approaches. One such method was the use of restriction fragment length polymorphism (RFLP) (Moller, et al 1985; Hongming, et al 1986; Inoko, et al 1986; Sheldon, et al 1986; Tilanus, et al 1986; Bell, et al 1987; Bidwell, et al 1987; Bodmer, J. et al 1987; McDaniel, et al 1987; Bidwell, et al 1988; Hyldig, et al 1988; Lin, et al 1989; Segurado, et al 1990; Dormoy, et al 1992). This method relies on the use of panels of restriction enzymes to digest DNA at specific sequence sites. The resulting digest is then run on a gel to separate the different sized DNA fragments. This is then Southern blotted onto a membrane and hybridised with a Class I or Class II specific labeled probe. Bands are then detected on an autoradiograph. The size and pattern of banding with different restriction enzymes and different HLA probes is used to determine the HLA specificity.

The size and pattern of bands identified by the HLA probes is a product of the specificity of the restriction enzyme. This in turn is determined by the polymorphism in nucleotide sequence present between different HLA specificities. The method therefore relies on such differences being at sites which can or cannot be cut by the different restriction enzymes. Analysis of the band patterns can be difficult, especially if the probe used is not particularly specific, or the enzyme does not define any useful polymorphism.

RFLP analysis of DNA has been useful in determining polymorphism at a time when very few HLA alleles had been sequenced. RFLP was not straightforward to perform. Each typing required a panel of enzymes which did not necessarily provide unequivocal determination of HLA specificity. The method was mainly applied to HLA Class II analysis with Class I RFLP analysis being considered difficult with no advantage over serology or IEF. Methods using the polymerase chain reaction (PCR), were being developed which allowed analysis of DNA to identify HLA specificities through known sequence polymorphism. The use of RFLP as it stood, was quickly negated by

these developments, as it provided no real information in terms of DNA sequence.

As discussed below, PCR allowed the specific amplification of regions of DNA for further analysis. To this end RFLP analysis has been applied to HLA PCR product (Maeda, et al 1989; Loiseau, et al 1991; Yunis, et al 1991; Hviid, et al 1992; Salazar, et al 1992; Tanaka, et al 1992; Lee, et al 1994; Medintz, et al 1994). The specificity of the PCR replaces the need for hybridisation with a HLA specific probe. Because many of the sequences of HLA alleles were known, enzyme restriction sites could be found within the PCR product at sites of polymorphism. This has allowed clearer interpretation of PCR-RFLP typing since it is within the context of known sequence. The combination of specific PCR with RFLP has been successfully used in several studies.

1.3.3 Polymerase Chain Reaction.

The polymerase chain reaction (PCR) has been an important and powerful development in the field of molecular biology (Saiki, et al 1985; Mullis, et al 1986; Mullis, et al 1987). It allows specific amplification of stretches of DNA sequence through repeated cycles consisting of DNA denaturation, annealing of specific primer to the DNA single strand followed by nucleotide extension from the primers using a thermally stable enzyme. PCR allows small amounts of DNA to generate substantial quantities of specifically amplified regions of DNA which can then be used in further analysis (Bell, 1989). Hence in tissue typing, the PCR reaction was used to specifically amplify the polymorphic regions of HLA genes. This HLA PCR product could then be analysed for its polymorphic differences which would establish the tissue type. Tissue typing by the two step approach of specific PCR amplification followed by detection of polymorphism has been described in a number of approaches. These include the use of PCR in combination with sequence specific oligonucleotide probing (PCR-SSOP) (Saiki, et al 1986), probing by reverse dot blot (Saiki, et al 1989), sequencing of the PCR product or sequence based typing (SBT) (Santamaria, et al 1992), heteroduplex analysis of PCR products (Clay, et al 1994), single stranded conformational polymorphism analysis of the PCR product (PCR-SSCP)

(Yoshida, et al 1992). Typing is also performed in a one-step PCR using sequence specific primers (Olerup, et al 1991). These approaches, used singly or in combination, have all been applied as DNA based methods for tissue-typing Class I and Class II HLA specificities.

1.3.4 Sequence specific oligonucleotide probing of PCR product (PCR-SSOP).

One initial and still popular method of DNA typing is the use of sequence specific oligonucleotide probes (SSOP) for detection of polymorphism within PCR amplified regions of HLA genes (figure 1.7). Initial PCR amplification of hypervariable regions of the HLA genes (e.g. exons 2 and 3 of the HLA-A, -B, -C genes and exon 2 of the DRB1 genes) provides a template which can be probed with labeled oligonucleotides which have specificity for particular sequence polymorphisms. Panels of these probes (SSOP) can then be used to identify which polymorphisms are present in the amplified DNA.

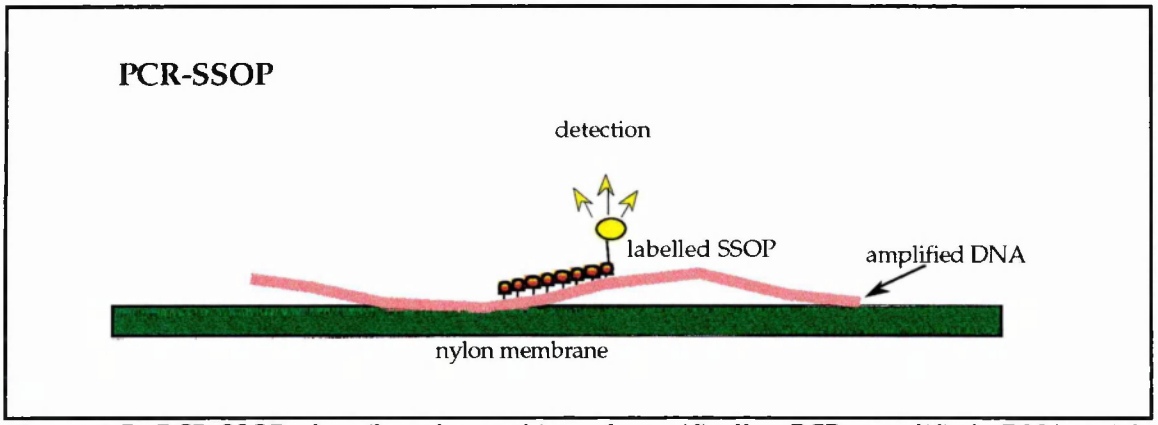


Figure 1.7. PCR-SSOP describes the probing of specifically PCR amplified DNA with labelled sequence specific oligonucleotides. PCR amplified DNA is fixed to a nylon membrane through UV cross-linking. Labelled probes are then hybridised with the DNA on the membrane. The membranes are then washed to an appropriate stringency, to remove any false hybridisation. Detection of hybridisation is through the labelled probe. With ^{32}P labelled radioactive SSOP, the hybridised membrane is exposed on an autoradiograph film. This is also true for chemiluminesence, in which the label can be reacted to emit light, which is detected on film. Chromogenic detection involves mainly biotin labelled probes which bind with a streptavidin/horseradish peroxidase ligand. the latter can then react with an appropriate substrate, causing a colour change to indicate hybridisation on the membrane.

Tissue typing by PCR-SSOP relies mainly on a pattern of hybridisation reactivity. This is to some extent hampered by the lack of polymorphism unique to a particular allele, and the more common presence of polymorphism shared amongst groups of specificities. What makes a particular HLA allele unique is the combination of these shared sequence polymorphisms along the length of the gene. The potentially complex patterns of detected reactivity may require detailed analysis to assign specificities, and certain heterozygous combinations may prove problematic to distinguish, as in most cases, both alleles at a HLA locus are simultaneously amplified in the PCR.

Interpretation of hybridisation patterns becomes more complex in relation to the lack of specificity within the PCR. A broadly specific PCR product requires an extensive panel of probes to differentiate between all the potential alleles amplified. This situation can be simplified through reducing the number of different alleles amplified, by introducing specificity into the PCR reaction.

Methods for PCR-SSOP typing therefore rely on a fairly specific PCR amplification which contains the hypervariable regions for probing. This may be a locus specific or even allele group specific amplification. This PCR product is then blotted onto a suitable membrane support and immobilised (UV crosslinked). Using an appropriate means of blotting, many samples can be loaded onto a single membrane. The membrane can then be probed with a labeled oligonucleotide under the required hybridisation conditions. Detection initially involved exposure of the membrane hybridised with radiolabeled probe onto an autoradiograph. To obtain a typing therefore requires multiple blotted membranes and multiple hybridisations, even though a membrane can be stripped of one probe and re-hybridised with another. Because typing required hybridisations with a panel of probes, a relatively large amount of radioactivity was used and so other non-radioactive methods of labeling have been considered.

A number of non-radioactive detection methods have now been developed. These include labeling the probe with biotin, which can then be detected by a colour change through use of a streptavidin/horseradish peroxidase conjugate (Bugawan, et al 1994). Other methods of probe labeling include the

use of digoxigenin-11-ddUTP (DIG), (Boehringer-Mannheim, Germany). Detection is through the use of an anti-DIG antibody conjugated to alkaline phosphatase. The DIG - anti-DIG complex can then be detected through the ability of the alkaline phosphatase to generate chemiluminesence with the appropriate substrate (Yoshida, et al 1992; Cereb, et al 1995).

PCR-SSOP Class II typing. PCR-SSOP was first applied to HLA Class II alleles, made possible through the availability of sequenced alleles. Initial studies concerned with typing HLA DQ specificities quickly spread to DR specificities as more sequence information became available. PCR-SSOP HLA Class II typing was applied to forensics studies (Higuchi, et al 1988) and disease association (Wordsworth, 1991; Hill, et al 1992) and histocompatibility matching for transplantation (Tiercy, et al 1989; Tiercy, et al 1991) Consolidation and standardisation of methods for PCR-SSOP typing happened through the 1991 International Histocompatibility Workshop (Kimura, et al 1992), which provided the forum for introduction of DNA based typing to histocompatibility laboratories worldwide. Participants in the workshop typed DRB1, DRB3, DRB4, DRB5, DQA1, DQB1, DPA1, DPB1 Class II loci using standard reagents and PCR-SSOP protocols. Around 120 probes were used in the study by about 132 laboratories. Each laboratory was expected to type around 100 individuals. The use of a large number of probes necessitated a computer program to interpret the complex hybridisation patterns possible through this typing. The success of Class II typing by this method and experience gained, provided the incentive for further improvements, particularly to detect through non-radioactive means and improve sample handling.

Although a method for non-radioactive detection was offered through the workshop (Kimura and Sasazuki, 1992), the majority of typing was done using ^{32}P labeled probes. The use of radiation was a problem for routine typing laboratories. Methods involving biotin labeled probes with subsequent colour detection (Erlich, et al., 1991a), or chemiluminesence using DIG labeling (Giorda, et al 1993; Chen, et al 1994; Cereb, et al 1995), have all been described and successfully incorporated into routine typing situations. PCR-SSOP has been described in microtitre plates. Each well contains a piece of DNA dotted membrane so allowing multiple simultaneous hybridisations and non-radioactive detection to be performed

across the plate (Garcia, et al 1995). Many histocompatibility laboratories have adopted the PCR-SSOP approaches described, influenced by difficulties in performing serological identification of Class II specificities and the resolution possible by DNA based methods.

PCR-SSOP Class I typing. HLA Class I PCR-SSOP has been developed in the wake of Class II. As already mentioned, this was partly because of the lack of sequence data for Class I alleles, but also because many workers were content with the continued use of serology. With sequence information becoming available over recent years, the application of DNA based typing methods to Class I typing became possible. The sequencing of increasing numbers of Class I alleles made clear that serology lacked the sensitivity to differentiate between certain specificities. For this reason, specificities such as the variants of HLA-A2, -A68 (Fernandez-Viña, et al 1992; Tiercy, et al 1994), or differentiating HLA-B35, and B53 (Allsopp, et al 1991; Hill, et al 1992) became the subject of Class I typing developments. With the realisation that PCR-SSOP could work for Class I, more general protocols started to emerge, describing broader typing for a particular locus. Hence PCR-SSOP methods have now been described for typing the HLA-A locus (Oh, et al 1993; Gao, et al 1994), the B locus (Yoshida, et al 1992; Fernandez-Viña, et al 1995; Middleton, et al 1995) and C locus (Levine, et al 1994).

Methods for typing have had to address the shared nature of sequence polymorphism found amongst HLA Class I alleles and the complex hybridisation patterns they potentially generate. The daunting complexity of interpreting hybridisation patterns obtained from probing a broadly specific Class I PCR product has made a level of specificity within PCR amplification a prerequisite for PCR-SSOP typing these alleles. As an example, a method for typing HLA-B uses five specific amplifications which divide the B locus alleles into manageable and interpretable groups for probing (Fernandez-Viña, et al 1995). The PCR product to be probed must include all the polymorphic sites necessary for determination of tissue-type. This usually entails amplification of exons 2 and 3 which code for the alpha-1 and -2 domains of the Class I molecule, using primer sites which specifically amplify the alleles of interest. Hence for the A locus, specific PCRs that amplifies all or most of exons 2 and 3 have been described (Oh, et al 1993; Allen, et al 1994).

At the B locus, the level of polymorphic complexity required added specificity in the PCR amplification. In an early method (Yoshida, et al 1992), the B locus was split into two PCR groups as determined by whether they contained the sequence coding for the Bw4 or Bw6 public epitopes. However, the region of the B locus alleles amplified, consisted of only half of exon 2, and many useful polymorphisms in exon 3 were excluded. This protocol has been improved by the inclusion of additional primer combinations that amplify across polymorphisms in exon 3 (Middleton, et al 1995). This allows many more polymorphisms to be tested, so enhancing the specificity and resolution of the method. A higher resolution system has also been described (Fernandez-Viña, et al 1995), in which an initial PCR screen determines which of 5 specific PCR reactions are used for probing. The appropriate PCR can then be performed and probed with the required set of oligonucleotides as required for the specificities amplified. The use of specific PCR which divides alleles into manageable groups for probing seems to be the route through which high resolution typing is achieved by PCR-SSOP. The combination of the described methods at the different loci, has allowed for general PCR-SSOP Class I DNA typing.

1.3.5 The reverse dot-blot approach to PCR-SSOP.

In parallel with development of the above conventional PCR-SSOP, a variation of the method known as the 'reverse dot blot' has also been described (Saiki, 1989). This is the reverse of the conventional PCR-SSOP approach since the oligonucleotide probes and not the amplified DNA, are immobilised onto the membrane (figure 1.8).

The labeled PCR product can then be simultaneously hybridised against all the probes bound on a single membrane. This innovation, available as a kit, has been widely used. It means that an individual can be typed quickly through one hybridisation event, rather than the lengthier conventional PCR-SSOP methods.

The raw data obtained by reverse dot blot is the same as conventional PCR-SSOP. It therefore may contain complex hybridisation patterns which require the aid of an appropriate computer program for their interpretation.

The format of the membrane, on which positives are identified by a colour change, can be quickly entered or even now scanned onto the computer which then determines the tissue type. This method is currently limited to Class II typing (Saiki, et al 1989; Erlich, et al 1991a) with Class I typing only recently being attempted (Bugawan, et al 1994). However, the success of conventional PCR-SSOP methods should mean in theory that the reverse dot blot approach should also be feasible for complete Class I typing, and this is currently under development.

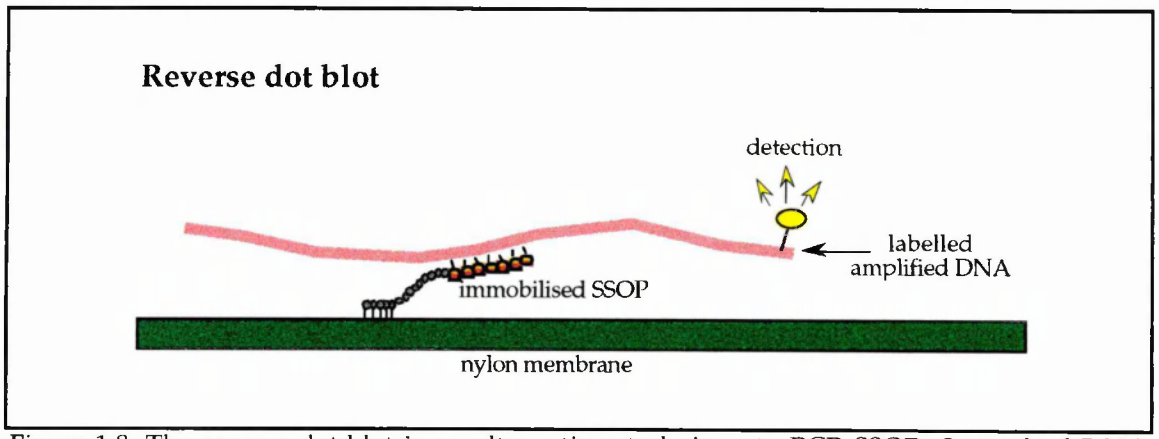


Figure 1.8. The reverse dot blot is an alternative technique to PCR-SSOP. Instead of DNA being fixed, the SSOP is immobilised onto the membrane. This has the advantage that a several different SSOP hybridisations can be performed simultaneously on the same membrane, hence allowing a sample to be typed in one go. Detection is through a labelled primer in the initial PCR reaction. The labelled PCR product, bound to the SSOP on the membrane, can then be detected usually through a chromogenic approach.

A variation of the reverse dot blot is immobilising probes in microtitre plates rather than nylon membranes. An advantage of using the microtitre plates is the array of dedicated equipment available for use with the format. In addition, detection of hybridisation can be performed in an appropriate plate reader. Detection in plates therefore makes semi or fully automated HLA typing possible. The use of microtitre plates to perform the reverse dot blot has been reported by a number of groups for typing HLA specificities (Giorda, et al 1993; Kawai, et al 1994).

1.3.6 PCR-SSP.

An alternative and subsequent method of typing to PCR-SSOP, is the use of sequence specific primers (SSP) in panels of PCR reactions (PCR-SSP), each reaction identifying a particular HLA specificity or group of specificities. As with the reverse dot blot, PCR-SSP offered an individual sample to be typed completely in one step, rather than the multiple hybridisations necessary for conventional SSOP procedures. In effect, PCR-SSP removes the need for probing a broadly specific PCR product by making the PCR amplification highly discriminatory. To achieve this level of specificity, SSPs were designed on the basis of the amplification refractory mutation system (ARMS) approach (Newton, et al 1989). This entails matching the 3' end of the primer with the sequence polymorphism it is required to detect. A mismatch at the 3' end of the SSP with the target sequence will under the correct conditions, inhibit amplification by disrupting the Taq enzymes ability to successfully read through and extend past the primer sequence (figure 1.9). The correct conditions rely on stringency, which is determined by the composition of the reaction mix and the annealing temperature in the PCR amplification program.

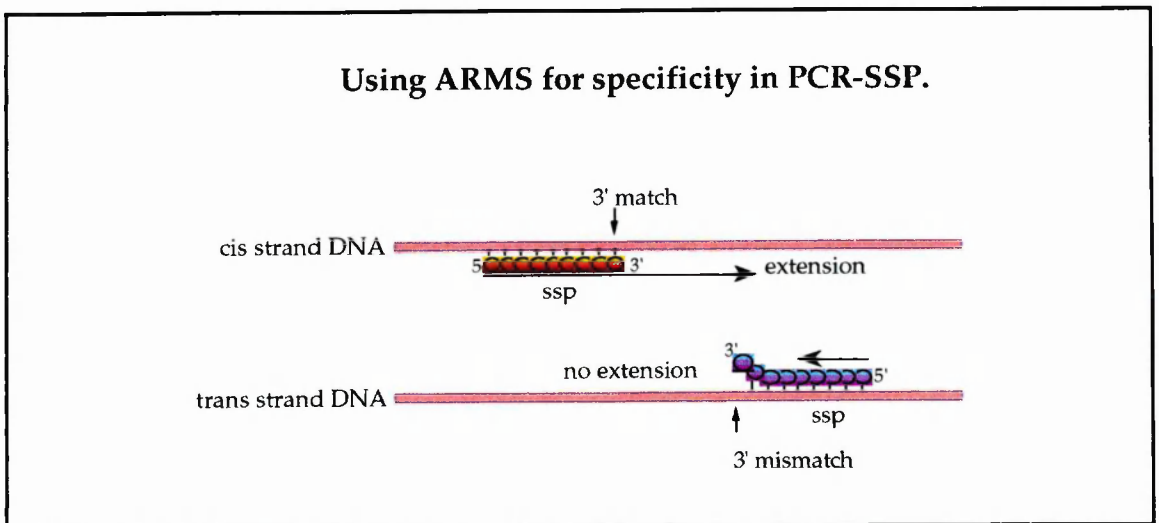


Figure 1.9. PCR-SSP requires highly specific PCR reactions to discriminate between different HLA types. This is achieved through ARMS, in which a 3' mismatch with the SSP inhibits amplification. For amplification to occur under the appropriate PCR stringency, both SSP's need to matched at their 3' residues for amplification to occur. Hence in the figure, although the SSP binding to the cis DNA strand is matched, and extension is possible, a 3' mismatch on the SSP binding to the trans strand inhibits extension, so causing the PCR reaction to fail.

All the PCR reactions in a typing panel are performed simultaneously under the same amplification conditions. This is achieved through design of the SSPs to have similar melting temperatures (Thein, et al 1986). This allows each PCR reaction in the typing panel to behave comparably under the similar levels of stringency. To identify the tissue type requires a panel of PCR reactions with each reaction specific for a particular HLA specificity. An internal control amplification is included in each PCR to avoid misinterpretation due to false negative PCR failure. Following the PCR, products are visualised on an agarose gel and the tissue type determined by the presence or absence of the appropriate bands in the appropriate lanes. The ease of interpretation of PCR-SSP typing is a major advantage of this approach. A more extensive description of the design of PCR-SSP reactions is given in chapter 3.2.2.

Multiplex PCR-SSP has also been investigated in which a panel of typing reactions is compressed, with more than one specific reaction present per reaction mix (Browning, et al 1994). Each individual SSP combination needs to produce a product which can be separated from its 'room-mates' on the basis of size. Further to this, all the SSPs used in a reaction mix must be compatible with each other to avoid inappropriate cross-priming and amplification in the presence of certain HLA sequences. This problem does have major implications since a newly identified HLA allele may interfere with the specificity of a multiplex typing panel and so require its redesign.

Class II PCR-SSP typing. The use of PCR-SSP for HLA typing, was first described for Class II specificities (Olerup and Zetterquist, 1991; Olerup and Zetterquist, 1992; Zetterquist, et al 1992; Olerup, et al 1993). An important study was the description of typing HLA-DR through analysis of sequence polymorphism located in exon 2 of the DRB1 alleles. Additionally, DR52 and DR53 were determined through analysis of DRB3 and DRB4 loci. The study demonstrated the ability to DR type using a panel of simultaneously performed PCR reactions run under the same amplification conditions. SSP identifying polymorphism at the 5' region of exon 2 are combined with SSP identifying polymorphism in the hypervariable region at the 3' end. The SSP combinations in the DRB typing paper produced product varying from 130-260 base pairs in length (Olerup and Zetterquist, 1992). The advantage of knowing the priming site of each SSP is that the resulting PCR product will

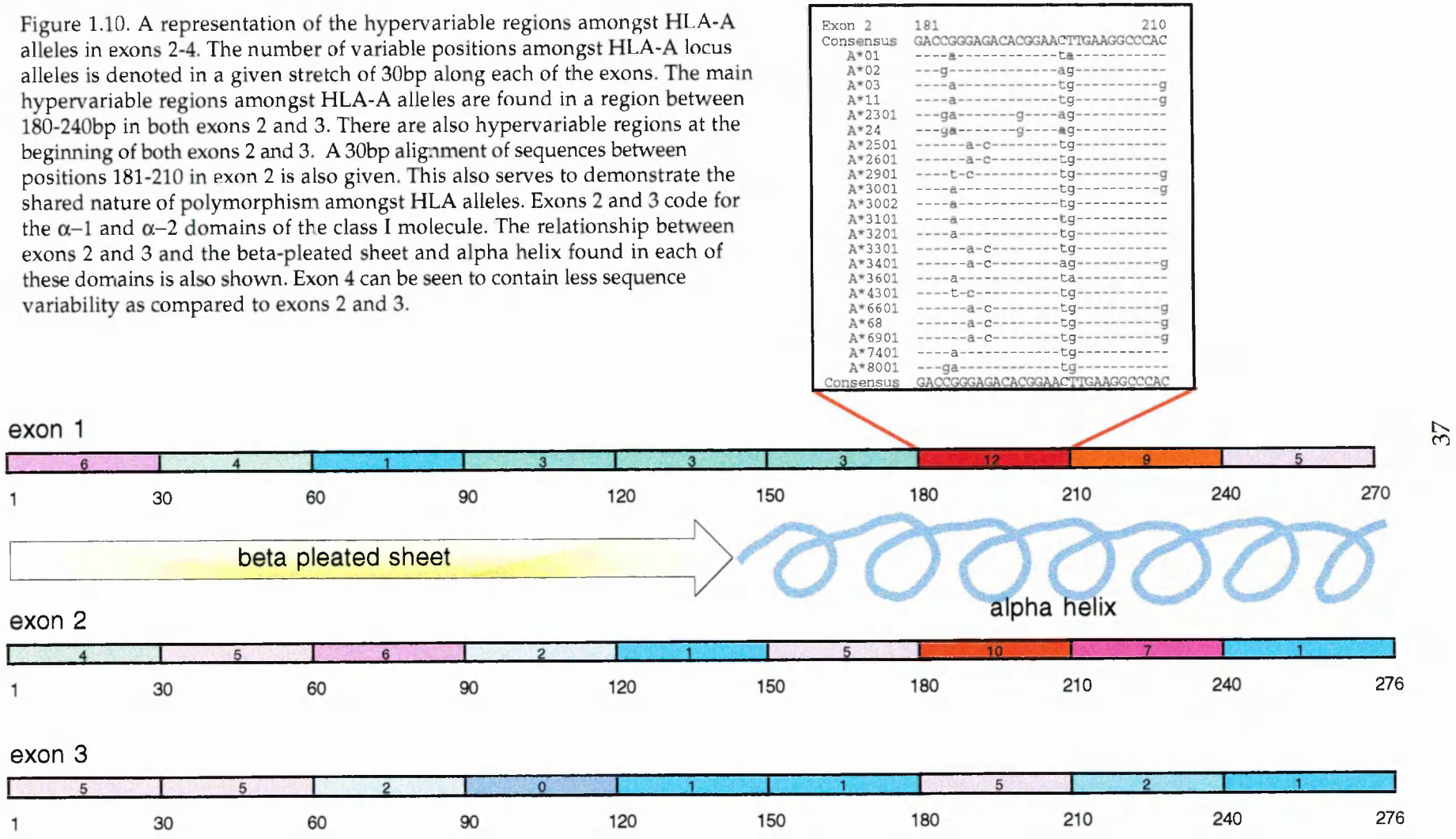
be of an expected size. This ensures that bands of an inappropriate size can be ignored. Development of PCR-SSP spread to typing at other Class II loci, namely DQ (Bunce, et al 1993; Olerup, et al 1993) and DP (Knipper, et al 1994) specificities.

Reports were made of complete DR typing, from blood to result, in 2 hours (Olerup, et al 1993). One perceived limitation of PCR-SSP is its non-suitability for large scale typing as compared to PCR-SSOP. This is mainly due to restriction in sample handling and equipment, especially if PCR reactions are performed in individual tubes. However, sample throughput has improved especially through the use of microtitre plates and the associated technology.

HLA Class I PCR-SSP typing. Identification of Class I alleles by PCR-SSP followed Class II through the restriction of available sequence data. Problems were envisaged due to the complexity of polymorphism defining Class I alleles as compared to Class II, especially since these are present mainly in two exons rather than one as in Class II typing. In someways the nature of polymorphism between HLA alleles suited PCR-SSP typing. The patchwork nature of polymorphism within hypervariable regions (figure 1.10), where sequence motifs may be shared between alleles of the same locus, other loci and pseudogenes, means that in many cases, more than one polymorphism has to be identified to ascertain the presence of an allele.

The PCR-SSP assay can to some extent address this problem, through the specificity of both SSPs in the PCR reaction. The main polymorphisms governing identification of these Class I specificities are found in exons 2 and 3 of the nucleotide sequence. Hence amplification was performed using one SSP in exon 2 in combined with another in exon 3, the product from which included the intervening intron when performed from genomic DNA. In certain cases, polymorphisms present in exons 1 and 4 have also been used for PCR-SSP typing (Bunce, et al 1994; Bunce, et al 1995a; Krausa, et al 1995a).

Figure 1.10. A representation of the hypervariable regions amongst HLA-A alleles in exons 2-4. The number of variable positions amongst HLA-A locus alleles is denoted in a given stretch of 30bp along each of the exons. The main hypervariable regions amongst HLA-A alleles are found in a region between 180-240bp in both exons 2 and 3. There are also hypervariable regions at the beginning of both exons 2 and 3. A 30bp alignment of sequences between positions 181-210 in exon 2 is also given. This also serves to demonstrate the shared nature of polymorphism amongst HLA alleles. Exons 2 and 3 code for the α–1 and α–2 domains of the class I molecule. The relationship between exons 2 and 3 and the beta-pleated sheet and alpha helix found in each of these domains is also shown. Exon 4 can be seen to contain less sequence variability as compared to exons 2 and 3.



The contribution of this thesis to the development of PCR-SSP typing at HLA-A locus will be discussed more extensively in the following chapters. Similar approaches were being applied to typing at HLA-C (Bunce, et al 1994). Identification of HLA-B alleles (Guttridge, et al 1994; Sadler, et al 1994) followed those described at the other loci, mainly due to the highly polymorphic and complex nature of this gene. Having demonstrated the feasibility of identifying Class I specificities, further improvement in resolution were made (Krausa, et al 1993; Bunce, et al 1994; Bunce, et al 1995a; Krausa, et al 1995a; Krausa, et al 1995b), showing that at each particular locus typing could well exceed that achievable through serology.

The advantages of PCR-SSP Class I typing over serology was really emphasised by the easy discrimination of HLA-C locus specificities. In performing HLA-C locus PCR-SSP typing, it was obvious that serology was subject to missing or misassigning HLA-C locus specificities. This has always been a weak-point for serology, due to the low expression of HLA-C on the cell surface, and crossreactivity of antisera. The ability to type for HLA-C locus alleles, including identification of certain C locus 'blanks' underlined the worth of DNA based typing.

HLA Phototyping. With improved resolution being shown also at the HLA-B locus (Bunce, et al 1995a), PCR-SSP typing had been adequately demonstrated at all the major HLA loci (Bunce, et al 1993; Olerup, et al 1993; Bunce, et al 1994; Bunce, et al 1995a; Krausa, et al 1996). It seemed a reasonable progression to combine all these published typing panels into a single integrated system. An obvious consequence of this was the problem of sample handling and performing the typing, if a large number of PCR amplifications have to be run simultaneously and loaded onto a gel for interpretation. The use of a 96 well microtitre plate for PCR provided a solution to this problem. The use of multitrack pipettes allowed samples to be quickly loaded onto a typing plate, and following amplification loaded onto a multitrack compatible gel apparatus. Moving away from reaction tubes to reaction plates greatly improved PCR-SSP as a technique and allowed a format, on which simultaneous typing at all the loci could be performed. The availability of a 192 well plate with compatible PCR machine has further improved the PCR-SSP typing procedure.

This multiple locus PCR-SSP typing procedure, known as 'phototyping' (Bunce, et al., 1995b) provides a complete DNA based alternative to serology, allowing Class I and Class II typing to be performed simultaneously. The method describes 144 PCR reactions covering HLA-A, -B, -C, DRB1, DRB3, DRB4, DRB5 and DQB loci (figure 1.11). DNA plus reaction mix are distributed across the typing plates preloaded with the primer mixes and, following amplification, the whole reaction panel visualised on one gel. Tissue type is then interpreted dependent on the pattern of reactivity with the reaction panel. As with any other DNA based typing approach, the publication of new sequences has to be scrutinised and the change in specificity of the reaction panel noted, with new or replacement reaction combinations added as appropriate.

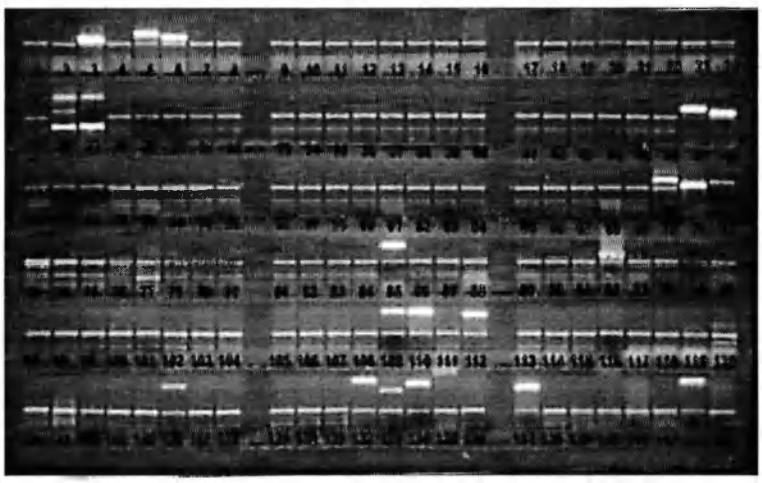


Figure 1.11. An example of an HLA PCR-SSP phototype. 144 PCR-SSP reactions are run simultaneously, encompassing HLA-A, -B, -C, -DRB1, -DRB3, -DRB4, -DRB5 and -DQB1 genes. Interpretation is reached through the presence or absence of the appropriately sized bands for each reaction following agarose gel electrophoresis. The common band present in each lane represents an internal control, which validates the PCR conditions. The sample types as A*03 (3) A*2301 (5,6) Bw4 (26) Bw6 (27) B*38 (47, 48) B*1518 (70, 71, 72) Cw*04 (85) Cw*0704 (92) DRB1*0301/4 (109, 110, 126) DRB1*04 (112) DRB3*0201 (132, 133) DRB4*01 (134) DQB1*02 (137) DQB1*0302 (143). The figures in brackets denote the reaction number defining the specificity. (Reproduced by permission of M. Bunce, Oxford)

The publication of a comprehensive PCR based system for defining Class I and Class II HLA specificities has prompted the view that DNA approaches

will soon replace serology as the main means for assigning tissue types (Dupont, 1995).

1.3.7 Heteroduplex analysis and conformational polymorphism.

Another approach to identifying HLA polymorphism, is through heteroduplex analyses (Clay, et al 1991; Rubocki, et al 1992; Sorrentino, et al 1992; Sorrentino, et al 1992; Sorrentino, et al 1993; Clay, et al 1994; Oka, et al 1994; Uhrberg, et al 1994). In the PCR, it is possible that single stranded DNA may re-anneal with an inappropriate strand to form a heterodimer or remain single stranded. Such events produce conformationally distinct structures which can be differentiated by their electrophoretic mobilities, and be useful in identifying HLA polymorphism.

Generation of PCR heteroduplexes have been used in matching HLA-DR specificities (Clay, et al 1991; Sorrentino, et al 1992) together with HLA-DQ (Rubocki, et al 1992), and HLA-DP (Sorrentino, et al 1992; Sorrentino, et al 1993; Clay, et al 1994). To enhance the discrimination offered by such analyses, an artificial single strand has been used, described as the universal heteroduplex generator (UHG) (Clay, et al 1994). This is a single strand of DNA containing engineered sequence polymorphisms in significant hypervariable regions. The PCR product can be denatured, reannealed with the UHG, with the resulting heteroduplex analysed in a more defined manner.

Heteroduplex analyses have obvious application for crossmatching in transplantation, although they do not identify the HLA type of donor and recipient. Mismatches can be determined by re-annealing donor and recipient amplified PCR product, specific for the particular required HLA region. Mismatches can then be detected through the differential migration of heteroduplex as compared to homoduplex PCR product through nondenaturing polyacrylamide gels.

The use of single stranded conformational polymorphism of PCR product (PCR-SSCP) is a further method used in typing. The electrophoretic mobility of single stranded DNA, under appropriate conditions, can be used to

discriminate between HLA alleles containing different polymorphisms. This approach has been applied to identifying HLA Class II (Shintaku, et al 1993) and HLA-B specificities (Yoshida, et al 1992).

1.3.8 Sequence based typing.

The full extent and complexity of HLA polymorphism can only be definitively explored through DNA sequencing. Even though methods such as PCR-SSOP and PCR-SSP offer high levels of resolution, they are limited within the context of testing for known sequence polymorphism. The use of heteroduplex or conformational polymorphism can only identify the existence of differences, not their exact nature in terms of sequence. Therefore the only definitive means of detecting such differences is DNA sequencing the region of interest, and identifying all the relevant polymorphisms. Although a great deal of polymorphism seems to be conserved and shared amongst several alleles at given hypervariable regions, unique polymorphic differences can exist, and these are not always within the confines of known hypervariable regions.

DNA sequencing (Sanger, et al 1977) can be a lengthy and labour intensive procedure, and in this form, unsuitable for routine tissue-typing. Innovations in DNA sequencing, particularly with the arrival of automated sequencing, has provided the means for rapid analysis of HLA polymorphism, and so making sequence based typing (SBT) a possibility.

Initial reports of SBT were for Class II (Santamaria, et al 1992), typing for DRB, DQA and DQB. Four PCR reactions are performed for DRB, one reaction for DQA1 and one reaction for DQB1. The four DRB PCR reactions can then sequenced with a general DRB primer. Three other sequencing reactions using more specific primers provide the sequencing information necessary for discrimination of DRB heterozygote combinations by preferential amplification of DRB1 over DRB3/4/5. For DQA and DQB each have a potential of two alleles per sequencing reaction. Although this approach showed the potential for HLA Class II typing for sequencing, it was still difficult to perform routinely. A similar approach was applied to Class I, in which two specific sequencing reactions were performed for each of the

HLA-A, -B and HLA-C loci (Santamaria, et al 1993). Class I and Class II typing could potentially be performed using SBT. Improvements in the method have already been made, particularly in using fluorophor labeled sequencing primers in conjunction with an automated sequencer (Versluis, et al 1993) and the use of magnetic beads as a sequencing solid support (Spurkland, et al 1993). Such approaches allow a larger throughput of samples, and potentially increase the generation of sequence information per reaction as compared to previous methods. Additionally, SBT has also been described for HLA-DP (Rozemuller, et al 1993).

Polymorphism not detected through methods such as PCR-SSOP and PCR-SSP, can be seen through SBT. Hence, previously unidentified alleles have been described (Santamaria, et al 1992; Santamaria, et al 1993), although the possibility of sequencing errors must also be considered (Domena, et al 1994). SBT provides the most information as compared to other DNA based typing methods. It does have limitations in the complexity of its performance and the requirement for expensive automated equipment. But as the facilities for doing such automated sequencing improve, SBT may soon be a routine technique for HLA typing and other important clinical indicators. The clinical relevance of typing at this resolution may well decide whether SBT is adopted as the method of choice.

1.3.9 Hybrid methods.

Many of the approaches described above can be used together in a complimentary fashion to elucidate the HLA type. The most common current combination of methods is to use serology for determination of Class I and a DNA based method for Class II. However, the recent advances and progress made in DNA based approaches, have greatly increased the choices available for tissue typing both Class I and Class II specificities. Additionally, different methods have been used together for determination of HLA type. Such developments reflect awareness of the capabilities of the different methods and how the specificity of one can provide the means for detection by another. Combining methods is usually indicative of the requirement for a high resolution of typing. Hence, using a highly specific PCR-SSP provides the material for analysis by SSOP, RFLP, SSCP,

heteroduplex or even another round of PCR-SSP as discussed in nested PCR-SSP. As an example, HLA-B alleles were divided by specific PCR into Bw4 or Bw6 products which were then hybridised by SSOP with additional information attained through SSCP analysis (Yoshida, et al 1992).

1.4 The context of this thesis.

This thesis presents some of the groundwork and developments made for DNA based approaches to Class I typing. This project was initiated at a time when Class I DNA typing was not seriously considered, and no real attempts had been made into this area. At this time, DNA typing was applied mainly to Class II specificities, and performed by PCR-SSOP.

With no available method for HLA Class I, this thesis aimed to develop a rapid, simple, easy to interpret, non-radioactive, highly definitive HLA-A locus typing system by PCR amplification of specific DNA sequences using sequence specific primers designed on the ARMS principle (Newton, et al 1989). The strategy taken in this project was to;

- First establish whether this approach was feasible in respect to identifying HLA-A locus specificities, through development of a very basic panel of typing reaction.

- Validate the basic reaction panel using DNA from serologically well characterised or sequenced cell lines where available.

- Increase the resolution to one equal or surpassing that of serology at the HLA-A locus.

- To develop strategies to achieve allelic definition at the HLA-A locus.

In addition, this thesis has also applied the methods developed in application to a number of population studies. Several new alleles have been identified and DNA sequenced through the course of this work. Additionally, a PCR protocol denoted 'gene mapping' was developed whereby the polymorphic composition of new alleles could be quickly

identified, prior to DNA sequencing. As a progression, the functional significance of small polymorphic sequence differences in a group of closely related HLA-A locus alleles was investigated, to underline the importance of high resolution typing. These areas of research are described over the following chapters.

Chapter 2

Materials & Methods.

2.1 The PCR reaction.

2.1.1. Preparation of DNA.

Through the duration of this project, different methods of DNA extraction were employed. Initially, DNA was prepared using standard proteinase K digestion followed by phenol/chloroform extraction and ethanol precipitation (Strauss, 1990). Latterly, DNA has been prepared with kits which negate the use of phenol. Two kits in particular have been used.

Kit 1. Nucleon II DNA extraction kit (Scotlab) uses SDS lysis of cells followed by sodium perchlorate protein precipitation. This is followed by chloroform separation, and addition of a silica slurry which binds any remaining protein in the aqueous phase. The silica can be spun out, and DNA ethanol extracted. The DNA pellet was resuspended in H_2O.

Kit 2. Puregene DNA Isolation kit (Flowgen) uses a lysis step followed by protein precipitation and isopropanol DNA precipitation. This kit does not require the use of chloroform and is simple to perform. This is the method currently used to prepare DNA for PCR-SSP typing.

A locus allele	Control cell line DNA in which allele identified.
A*0101	FMB, KLO, KAS011, WIN, VOO, **MOLT4**
A*0102	**DAUDI**
A*0201	**JY**, KT2, GRC150, JAP.NF, WDV, C073, RSA.ND, CALOGERO, OMW, DOP.ND, W#7079, **GRC138**
A*0202	**M7**
A*0203	**DK1**, GEE018
A*0204	**RML**
A*0205	**WT49, AM**
A*0206	AT, **CLA**, KY, W#7079
A*0207	**KNE**, AP
A*0208	**KLO**
A*0209	**OZB**
A*0210	**XLI.ND**
A*0211	**KIME, GRC138**, GRC212
A*0212	**KRC033**
A*0213	**SLUGEO**
A*0214	**1S**
A*0215N	-
A*0216	**TUBO**
A*0217	**AMALA**
A*0301	VEN, M7, JS, VOO, PLH, SNA,
A*0302	CGM1
A*1101	LATIF, WEWAK, APA, SLUGEO
A*1102	-
A*2301	WT51
A*2402	WEWAK, AKIBA, KT3, CAM020, ZM
A*2403	THA1742, **APA**
A*2404	-
A*2405	-
A*2406	-
A*2501	MOLT4, **BM92**
A*2601	PA, QBL, JAP.NF
A*2602	-
A*2603	-
A*2604	-
A*2901	GEE018
A*2902	C073, PITOUT, CGM1, 32/32, SWEIGH007,
A*3001	**LBF, RSH**, XLI.ND, 35020
A*3002	32/32, **CRB**
A*3003	**JS**
A*3004	**W7(CC)**
A*3101	GRC150, VEN, DAN723, LATIF, KRC103
A*3201	FMB, **AM**
A*3301	**LWAGS**, IBW9
A*3302	-
A*3303	**HOR**, AP, DOP.ND, DK1, 32367, THA1742
A*3401	RSA.ND
A*3402	*#11, 1528, 1337*
A*3601	*2D, #25*
A*4301	PSLD
A*6601	**TEM**, IDF, DYC.ND
A*6602	**CRB**
A*6801	**LBS**, GRC212
A*6802	SNA, **PA**, AMAI, RSH
A*6901	**IDF, ZM**
A*7401	32367, EEK.ND
A*8001	**35020**

Table 2.1. List of control cell line DNA used to test HLA-A locus alleles. Cell lines in bold have been sequenced for that particular allele. Cell lines in italics are local samples. Alleles not identified denoted by (-).

In developing the method, DNA was prepared mainly from cultured EBV transformed B cell lines, but also from blood samples. Blood was collected into EDTA rather than heparin, which is reported to inhibit amplification in the PCR reaction (Satsangi, et al 1994). Samples were obtained from a variety of sources, and these are described in the relevant sections of the thesis. Many of the HLA characterised and sequenced DNA samples were obtained from the Tissue Antigen Laboratory, ICRF, London. A list of some of the samples is given in Table 2.1

2.1.2. Components of the PCR reaction.

The basic components of the PCR reactions used in this thesis, are listed below. These have generally remained consistent thoughout the course of the project. Cycling parameters in terms of incubation times and temperatures have varied and evolved, mainly in response to changes in PCR machines used and the necessity to attain correct stringency as the resolution of the method increased.

10X PCR Buffer;	1095mg Ammonium sulphate (166mM)
	33.5ml 1M Tris pH8.4 (670mM)
	6.7ml 0.5M EDTA pH8.0 (67mM)
	85mg BSA (0.17%)
	dH_2O to 50ml

Sterilise through 0.22mm filter - store in 1ml aliquots at -20^{o}C.

PCR mix;	10X buffer	3000µl	
	dNTPs (100mM)	60µl	x4
	$MgCl_2$ (100mM)	600µl	
	Internal control primers (20mM)	300µl	x2
	dH_2O	19600µl	

Dispense into 1ml aliquots and store at -20^{o}C. To a 1ml aliquot, add 62.5µl of each SSP (4µM) and 5µl Taq (5U/µl). This mix can be dispensed into tubes or plates, at the appropriate volumes. The mix should be overlaid with mineral oil and stored frozen at -20°C. This mix constitutes 90% of the PCR reaction leaving 10% in which the test DNA can be added.

Approximately 100-200ng genomic DNA is added per 50µl PCR reaction. This needs to be clean and of good quality.

Making up larger batches of buffer/PCR mix and aliquoting these into usable volumes, reduces the amount of variation between typings and leads to more consistent results.

Storing 'ready to use' panels frozen leads to rapid typing as outlined above. Such prepared panels can be stored frozen at -20°C. Successful amplification has been noted with reactions stored frozen in excess of 6 months.

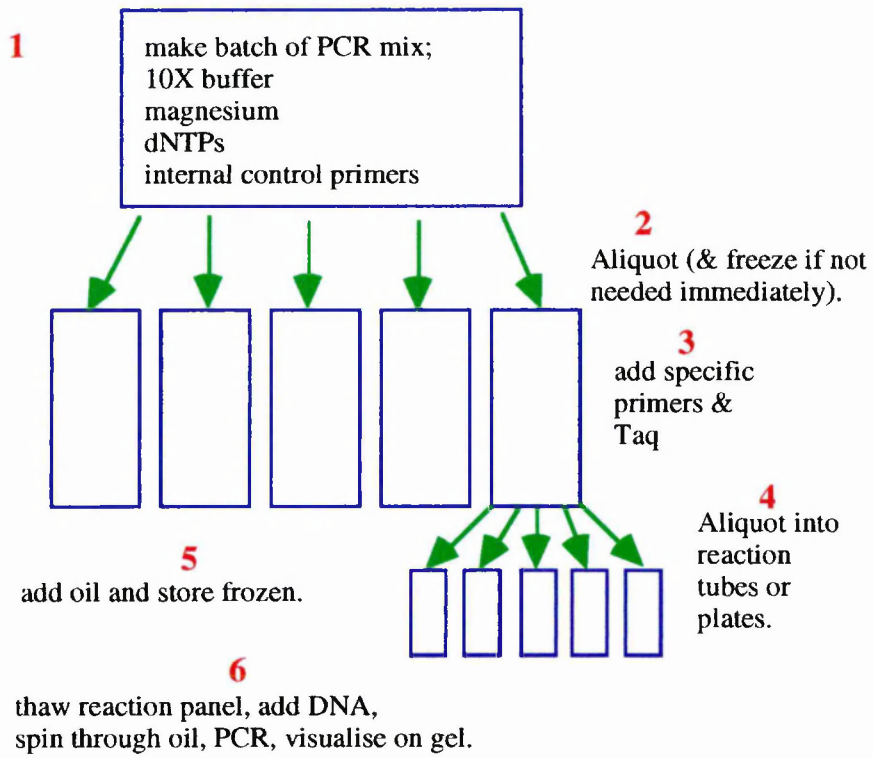


Figure 2.1. Schematic representation of preparation of PCR reaction mixes and typing panels.

Coding SSP	SSP sequence (5'-3')	SSP position (5'-3')	
		exon	nucleotide location
CLI5'	CCC ACT CCA TGA GGT ATT TC	2	4-23
AL#1	ACC AGG AGA CAC GGA ATA	2	181-198
AL#2	ACA GAC TGA CCG AGC GAA	2	212-229
AL#3	GAC GGG GAG ACA CGG AAA	2	180-197
AL#4	GAG TAT TGG GAC CGG AAC	2	171-188
AL#5	CAC TCC ATG AGG TAT TTC TC	2	6-25
AL#6	CGG AAT GTG AAG GCC CAG	2	192-209
AL#7	GCG ACG CCG CGA GCC A	2	112-127
AL#8	GCC GGA GTA TTG GGA CGA	2	167-184
AL#9	TAT TGG GAC GAG GAG ACA G	2	174-192
AL#10	GAT AGA GCA GGA GAG GCC T	2	152-170
AL#11	CAC AGA CTG ACC GAG AGA G	2	211-229
AL#12	CCG GCC CGG CAG TGG A	2	38-53
AL#13	TGG ATA GAG CAG GAG GGT	2	150-167
AL#14	AGG CCC ACT CAC AGA CTC	2	202-219
AL#15	CCT GCG GAT CGC GCT CC	2	230-246
AL#16	GAC CAG GAG ACA CGG AAT A	2	180-198
AL#17	CGG AGT ATT GGG ACC TGC	2	169-186
AL#18	GAC CGG AAC ACA CGG AAA	2	180-197
AL#19	CAG CGA CGC CGC GAG C	2	110-125
AL#20	TGA AGG CCC ACT CAC AGA T	2	199-217
AL#21	ACT ACA ACC AGA GCG AGG A	2	250-268
AL#22	CAC TCC ATG AGG TAT TTC TT	2	6-25
AL#23	GGC CCG GCC GCG GGA	2	40-54
AL#24	CAC GCA GTT CGT GCG GTT T	2	89-107
AL#25	GAG GTA TTT CTA CAC CTC CA	2	14-33
AL#26	GAC GGG GAG ACA CGG AAT	2	180-197
AL#27	CAC TCC ATG AGG TAT TTC TA	2	6-25
AL#28	GCG ACG CCG CGA GCC G	2	112-127
AL#29	CAG CTC AGA CCA CCA AGC A	3	145-163
AL#30	CGG AAT GTG AAG GCC CAC T	2	192-210
AL#31	CGG AAT GTG AAG GCC CAG T	2	192-210
AL#32	CGA GTG GAC CTG GGG AC	2	222-238
AL#33	GAC CTG GGG ACC CTG CG	2	228-244
AL#34	CTC ACA GAC TGA CCG AGC	2	209-226
AL#35	GGA TGG AGC CGC GGG CA	2	130-146
AL#36	ATA GAG CAG GAG GGG CCG	2	153-170
AL#37	CCT CGT CCC CAG GCT CT	2	(11)-5
AL#38	AGC CCC GCT TCA TCG CC	2	55-71
AL#39	CCG TGG ATA GAG CAG GAG A	2	147-165
AL#40	GTC CGG AGT ATT GGG ACG	2	166-183
AL#41	CAC TCC ATG AGG TAT TTC AC	2	6-25
AL#42	GAC GCC GCG AGC CAG AA	2	114-130
AL#43	TGT CCC GGC CCG GCA GT	2	34-50
AL#44	CAC AGA CTG ACC GAG TGG	2	211-228
AL#45	TCA CAG ACT GAC CGA GAG A	2	210-228
AL#46	GAG CCC CGC TTC ATC GCA	2	54-71
AL#47	CTA CAA CCA GAG CGA GGC	2	251-268
AL#48	CTG CTA CTC TCG GGG GCT	1	31-48
AL#49	GAC CAG GAG ACA CGG AAT G	2	180-198
AL#50	CGA ACC CTC CTC CTG CTA	1	19-36
AL#51	TCC ATG AGG TAT TTC TTC ACA	2	9-29
AL#53	CGA CTC GCA GTT CGT GCA	2	86-103
AL#54	AAG GCC CAC TCA CAG ACT A	2	201-219
AL#55	AAG GCC CAC TCA CAG ATT G	2	201-219
AL#56	GAC AGC GAC GCC GCG ACC CG	2	108-127

Table 2.2 Coding strand Sequence specific primers and locations.

Non-coding SSP	SSP sequence (5'-3')	SSP position (5'-3') exon	nucleotide location
AL#A	CTC CAG GTA GAC TCT CCG	3	212-195
AL#B	CGT CGT AGG CGT ACT GGT	3	87-70
AL#C	TGT AAT CCT TGC CGT CGT AA	3	99-80
AL#D	CAC TCC ACG CAC GTG CCA	3	229-212
AL#F	GCG CAG GTC CTC GTT CAA	3	122-105
AL#G	CCG TCG TAG GCG TGC TGT	3	88-71
AL#H	CAA GAG CGC AGG TCC TCT	3	127-110
AL#I	TCT CTG CTG CTC CGC CG	3	200-184
AL#J	TCA ACT GCT CCG CCA CAC	3	198-181
AL#K	TCT CAA CTG CTC CGC CCA	3	200-183
AL#L	CAA GAG CGC AGG TCC TCG	3	127-110
AL#M	CAG CCA AAC ATC ATC TGG AG	3	31-12
AL#N	CCC CAC GTC GCA GCC AA	3	41-25
AL#O	GTC GTA GGC GTG CTG TTC A	3	86-68
AL#P	GCC TCC CAC TTG CGC TG	3	175-159
AL#Q	CTC CAG GTA GGC TCT CAA	3	212-195
AL#R	CTC CAG GTA GGC TCT CTG	3	212-195
AL#S	CGC CTC CCA CTT GCG CTT	3	176-159
AL#T	CTC AAC TGC TCC GCC CG	3	199-183
AL#U	CTC CAG GTA GGC TCT CC	3	212-196
AL#V	GAG CCA CTC CAC GCA CGT	3	233-216
AL#W	AGG TAT CTG CGG AGC CAC	3	244-227
AL#X	GCC CGT CCA CGC ACC G	3	231-216
AL#Y	CCG CGG AGG AAG CGC CA	3	64-48
AL#Z	GGT ATC TGC GGA GCC CG	3	243-227
AL#AA	CTC TCC ACT GCT CCG CCT	3	201-184
AL#AB	CAC GTC GCA GCC ATA CAT TA	3	38-19
AL#AC	AGC CAC TCC ACG CAC CG	3	232-216
AL#AD	GAG CCA CTC CAC GCA CTC	3	233-216
AL#AE	TCC GCC TCA TGG GCC GT	3	190-174
AL#AF	ACG TCG CAG CCA TAC ATC A	3	37-19
AL#AG	CTT CCC GTT CTC CAG GTG	3	257-240
AL#AH	CAG GTA GGC TCT CAA CTC G	3	209-191
AL#AI	TGT CCG CCG CGG TCC AA	3	141-125
AL#AJ	CCC ACT TGT GCT TGG TGG	3	171-154
AL#AK	ACT GGT GGT ACC CGC GC	3	77-59
AL#AL	TGG AAG GTT CCA TCC CCT T	4	178-160
AL#AM	CGT CGT AGG CGT CCT GCC	3	87-70
AL#AN	CTC TCA ACT GCT CCG CCT	3	201-184
AL#AO	CTC TCT GCT GCT CCG CCA	3	201-184
AL#AP	CAT CCA GGT AGG CTC TCA A	3	213-195
AL#AQ	CCA TCC AGG TAG GCT CTC T	3	214-196
AL#AR	TGG TAC CCG CGG AGG AG	3	70-54
AL#AS	CCG TCC GAC CCC ACG TC	3	49-33
AL#AT	CTG CTC CGC CGC ATG GA	3	194-178
AL#AU	GTG CTT GGT GGT CTG AGC T	3	164-146
AL#AV	CTC CAG GTA GGC TCT CCA	3	212-195
AL#AW	TGG CCC CTG GTA CCC GT	3	(14)-273
AL#AX	TCC ACT GCT CCG CCA CAT	3	198-181
AL#AY	CTC CAC TGC TCC GCC AC	3	199-183
AL#AZ	GGC CGC CTC CCA CTT GT	3	179-163
AL#BA	TAG AAA TAC CTC ATG GAG TGA	3	25-5
AL#BB	GAA GCG GGG CTC TCC ACT	2	65-48
AL#BC	ATG AAG CGG GGC TCT CCA	2	67-50
AL#BD	GGT CCC AAT ACT CAG GCC T	2	183-165
AL#BE	AAG CGG GGC TCT CCA CT	2	64-48
AL#BF	CCC CAC GTC GCA GCC AT	3	41-25
AL#BG	CGT CGC AGC CAT ACA TCC	3	36-19
AL#BJ	CCG ACC CCA CGT CGC AGG CAC (mm)	3	45-25
AL#BK	AGC CCG TCC ACG CAC TC	3	232-216

Table 2.3 Non-coding strand Sequence specific primers and locations.

Non-coding SSP	SSP sequence (5'-3')	SSP position (5'-3') exon	nucleotide location
AL#BL	CTC TCT GCT GCT CCG CCT	3	201-184
AL#BM	CCG ACC CCA CGT CGC AGG CAT	3	45-25
AL#BN	CCT TGC CGT CGT AAG CGT T	3	93-75
AL#BO	GAG CCA CTC CAC GCA CAG	3	233-216
AL#BP	CAG CCA AAC ATC ATC TGG AG	3	31-12
AL#BQ	ACG TCG CAG CCA AAC ATC A	3	37-19
AL#BR	AGC CAT ACA TCC TCT GGA G	3	30-12
AL#BS	ATG CTG CAC ATG GCA GGT T	4	242-224
AL#BT	AGC CAT ACA TCC TCT GGA C	3	30-12
2901	CCG TCC GAC CCC ACG TG	3	49-33
A0207	CCC CAC GTC GCA GCC AC	3	41-25
CLI3'	CCT TCC CGT TCT CCA GGT	3	258-241

Table 2.3 continued. Non-coding strand Sequence specific primers.

Coding primer (B2M#1)	5'-CGA TAT TCC TCA GGT ACT-3'
Non coding primer (B2M#2)	3'-CAA CTT TCA GCA GCT TAC-3'

Table 2.4 Internal control primers (B2M#1+B2M#2) amplifying a 330bp region of the β-2microglobulin gene.

Sequence specific primers (SSP) were synthesised by the Imperial Cancer Research Fund, Oligonucleotide Synthesis Service, Clare Hall Laboratories, UK and by R&D Systems Europe Ltd, Abingdon, UK. SSP were either provided as quantified precipitates or solutions. These were then diluted to either 100μM or 20μM stocks in sterile distilled water and frozen at -20°C. SSP working stocks at 4μM were also made and stored frozen at -20°C. Each PCR reaction comprised a pair SSP's, one to the coding strand and the other to the non-coding strand. A list of the SSP sequences, and their anealing positions within the HLA-A gene sequence is given in Tables 2.2 and 2.3. The location of each primer is given as the 5' to 3' nucleotide postion in the given exon. The internal control primers are listed in Table 2.4

The thermal cyclers used include a Hybaid Thermal Reactor in which PCR amplification was performed in 0.6ml eppendorf tubes, a MJ-Research PTC-100 using non skirted 96 thermowell plates (Costar), and a MJ-Research PTC-200 using either non skirted 96 thermowell plates or 192 thermowell plates (Costar). The PCR conditions and cycling conditions are described in the relevant sections.

2.1.3. Agarose gel electrophoresis.

10µl of PCR products from the typing reactions were mixed with 2µl of bromophenol blue/xylene cyanole loading buffer prior to loading onto 2% agarose gel made in 0.5X TBE buffer prestained with ethidium bromide.

Initially, when the number of reactions were small and performed in 0.5ml eppendorf tubes, electrophoresis was performed in Biorad Mini Sub™ DNA cells. Expansion in the number of typing reactions required that PCRs were performed in 96 or 192 well microtitre plates. This required the use of Flowgen gel apparatus compatible with the use of multitrack pipettes. This gel apparatus allows screening of PCR products from 3x96 well plates.

Loading buffer;	0.2% bromophenol blue 0.2% xylene cyanole 50% glycerol in H_2O
Gels;	2% agarose in 1/2X TBE Ethidium Bromide (3µl 10mg/ml per 50ml gel)

(1L 10X TBE = 108g Tris base + 54g Boric acid + 40ml 0.5M EDTA)

PCR products run out on the gel were then visualised on a UV transilluminator and an image recorded using a polaroid camera.

2.1.4. PCR-SSP Phototyping.

The culmination of developments by a number of workers, including the contribution of this thesis, has allowed a general PCR-SSP system to be described (Bunce, M. et al., 1995), which identifies specificities at HLA-A, -B, -C, -DRB1, DRB3, DRB4, DRB5 and DQB1 loci. In brief, 26µl of DNA (200-300µg/ml) was added to an aliquot of reaction mix containing all the PCR components (PCR buffer, dNTPs, $MgCl_2$, Taq polymerase) with the exception of SSP primer pairs. The 144 SSP primer pair mixes were pre-aliquoted under oil in 96 well microtitre plate, and stored frozen at -20°C, ready for use. Each specific reaction also contained an internal control reaction. The

DNA and reaction mix were distributed across the typing panel and PCR performed through 30 cycles on an MJ-Research PTC-200 Thermal Cycler. The Cycling parameters on this thermal cycler comprised;

(96°C-50s, 70°C-50s, 72°C-50s)5 cycles +
(96°C-50s, 63°C-50s, 72°C-50s)21 cycles +
(96°C-50s, 55°C-50s, 72°C-50s)4 cycles

PCR products were then analysed through electrophoresis on an ethidium bromide stained agarose gel. Loading of samples onto the gel was facilitated by use of multitrack pipettes to which a dedicated gel apparatus (Flowgen, UK) was compatible. Using a UV transilluminator, a Polaroid image of the gel was taken and the HLA type of the test samples was determined by the presence or absence of PCR products of expected size in the appropriate wells of the reaction panel.

2.1.5. List of materials used for PCR.

Reagent	Company & Cat. No.
Nucleon II DNA extraction kit	Scotlab
Puregene™ DNA Isolation kit	Flowgen D-5000
Ammonium Sulphate	Sigma A-4418
Magnesium Chloride	Sigma M-9272
Bovine Serum Albumin (BSA)	Sigma A-6793
Ultrapure dNTP Set	Pharmacia 27-2035-01
Trizma Base	Sigma T-1503
EDTA disodium salt	BDH 10093
Agarose electrophoresis grade	GibcoBRL 15510-027
Boric Acid	Sigma B-6768
Ethidium Bromide	Sigma E-8751
Bromophenol Blue	Sigma B-6131
Xylene Cyanole FF	Sigma X-4126
Glycerol	BDH 10118
Ethanol	Hayman Ltd SIN1170
Isopropanol	Fisons P/7500/PB17
Phenol	Rathburn RP3024
Chloroform	Fisons C/4960/PB17
Proteinase K	Sigma P-2308
Mineral oil	Sigma M-3516

2.2 DNA cloning and sequencing.

2.2.1. cDNA preparation.

The general protocol adopted was to extract mRNA (mRNA Isolation kit, Stratagene) from the cell line of interest, prepare cDNA by RT-PCR (Superscript™ first strand cDNA synthesis kit, GibcoBRL), and PCR amplify out the genes of interest using HLA-A locus specific primers. The 100µl PCR mix consisted of;

10µl - 10X PCR buffer (as above)
21µl - cDNA (as prepared with above cDNA kit)
10µl - primer mix (2µM of each primer given in figure 2.2)
1µl - Taq enzyme at 5U/µl
58µl - distilled sterile water
overlayed with 15µl of mineral oil

(dNTP's, $MgCl_2$ are present in the cDNA fraction of the PCR mix).

The cycling parameters as performed in 0.6ml eppendorf tubes in a Hybaid thermal reactor were;

(95°C-1min, 65°C-1min, 72°C-2min)5cycles +
(95°C-1min, 55°C-1min, 72°C-2min)25cycles

HLA-A5'EcoR1 5'-gTC.gAA.TTC.gCC.gCC.ATg.gCC.gTC.ATg.gCg.CCC.CgA-3'
EcoR1 KOZAK START

HLA-A3'BamH1 5'-CgC.ggA.TCC.TCA.TCA.CAC.TTT.ACA.AgC.TgT-3'
BamH1 STOP

Figure 2.2. Primer sequences used for amplification of the full length HLA-A locus gene.

The PCR product was then run out on an ethidium bromide stained TAE Agarose gel, and the 1100bp full length HLA-A locus PCR product cut out

and purified from the gel (Easiclene DNA Isolation kit, Scotlab Bioscience).

The HLA-A specific primers which amplify the full length gene, incorporate EcoR1 and BamH1 restriction sites to facilitate the correct orientation of the HLA gene when ligated into the appropriate vector. The primers also incorporate kozak, start and stop codons which assist in the expression of the HLA-A gene, within the appropriate vector.

2.2.2. cDNA cloning

The PCR product potentially contains two HLA-A locus specificities, and therefore a cloning step is necessary to isolate the allele of interest. To this end, the isolated PCR product was either digested with EcoR1/BamH1 restriction enzymes (Promega) and ligated into a similarly digested M13-mp18 and 19 bacteriophage vector for single stranded DNA sequencing, or ligated directly into pMOSblue T-vector (Amersham Life Sciences), for double stranded sequencing.

The ligation reaction used typically consists of;

1µl - 10X Ligase buffer
0.5µl - 100mM DTT
0.5µl - 10mM ATP
1µl - cut vector
2µl - insert (cut as approriate)
5µl - distilled sterile water
0.5µl - T4 DNA ligase 4U/µl.

The above ligation reaction was incubated at 16°C overnight.

The main reagents for ligation and transformation were provided in the pMOSblue T vector kit (Amersham). The competent cells provided were transformed with the insert ligated bacteriophage vector. This required addition of 1µl of the ligation mix to 20µl of chilled competent cells. This was incubated on ice for 30 minutes prior to a heat shock at 42°C for 40 seconds. The mixture was then rested on ice for 2 minutes prior to addition of 80µl of SOC media and incubation at 37°C for 1 hour. The cells were spread onto L agar plates containing 50µg/ml ampicillin (and 15µg/ml

tetracyline for the pMOSblue T-vector transformations). White/blue selection was used to screen for and exclude 'blue' recombinants by including 50µl of 50mg/ml X-gal+100mM IPTG in each L agar plate. Plates were incubated at 37°C overnight, and clear plaques (M13) or white colonies (pMOSblue) were picked and cultured for around 6-8hrs in 3ml of L Broth. 1ml of this culture was frozen at -20°C in Hogness Media, as a stock for further culturing of the plasmid bearing cells.

(i). Single stranded M13mp18/19 small-scale DNA preparation.

From 2ml of cells in L Broth, 200µl of PEG/NaCl was added to 1.2ml of culture supernatant. This mixture was vortexed and allowed to stand for 15 minutes prior to centrifugation at 14,000rpm in an eppendorf benchtop centrifuge. All the supernatant was removed and the pellet was resuspended in 100µl TE. 50µl of equilibrated Phenol was added to the resuspended pellet, mixed and centrifuged. The aqueous phase was then removed and added to an equal volume of chloroform and centrifuged. The aqueous phase was then added to a fresh tube containing 300µl of a 25:1 mixture of ethanol to 3M Na Acetate. This was then mixed and left to stand for 15 minutes. The precipitated single stranded DNA was recovered by centrifugation for 10 minutes at 4°C, and the supernatant removed. The pellet was then washed in 70% ethanol, allowed to air dry for 10 minutes prior to resupension in 50µl TE.

(ii). Double stranded pMOSblue T-vector small-scale DNA preparation.

Plasmid DNA was prepared from 2ml of cells cultured in L Broth. DNA was prepared using the QIAprep Spin Plasmid Kit (QIAGEN), using their standard protocol. Around 20µg of plasmid DNA was recovered in a volume of 100µl H_2O.

Prior to sequencing, this double stranded plasmid preparation required denaturation. This was achieved through the addition of 0.1 volumes of 2M NaOH/2mM EDTA to the plasmid DNA which was then incubated at 37°C for 30 minutes. To neutralise the mixture, 0.1 volumes of 3M Na Acetate (pH5.5) was added and DNA is precipitated through addition of 3 volumes of absolute ethanol and left at -75°C for 15 minutes. The DNA pellet was

washed in ethanol and resuspended in the appropriate volume of distilled H_2O for sequencing.

2.2.3. DNA sequencing reaction protocol.

DNA sequencing of the prepared plasmid DNA was performed using standard chain termination reactions (Sanger, F. et al., 1977), as facilitated in the use of the Sequenase™ version 2.0 DNA sequencing kit (USB - Amersham Life Sciences). In brief, this involved;

- Initial annealing of 1pMol of primer with 3-5µg plasmid DNA in the presence of reaction buffer. This was incubated at 65°C for 2 minutes and allowed to cool slowly.

- For the labelling reaction, DTT, labeling mix, (α-^{35}S)dATP and sequenase were added to the annealing mixture.

- The termination reactions were performed by addition of the labelling reaction to each of the four 37°C prewarmed termination mixes (ddATP, ddCTP, ddGTP, ddTTP). After 5 minutes, the reactions were stopped by addition of the stop solution. The sequencing reactions were heated for 2 minutes at 80°C prior to loading onto a gel.

- For each sample, 2.5µl of each sequencing reaction (A,C,G,T) was loaded onto a 5% acrylamide/bis (19:1) Sequagel™ sequencing gel (National Diagnostics) using a Biorad Sequi-gen™ vertical gel system and 3000xi powerpack. The sequencing reactions were run through the pre-warmed (50°C) gel at @1750V; 50W for approximately 3-4hrs. The gel was then placed onto a paper backing and dried prior to exposure with Kodak X-omat film overnight or longer as required. The film was developed and the sequence read.

2.3. Transfection of HLA genes into CIR cells.

2.3.1. Cloning of expression vector.

As above, full length HLA-A cDNA was prepared and ligated into the

pMOSblue T-vector, cloned and sequenced to verify the integrity of the HLA insert. To transfect into the Class I reduced (CIR) cell line, the HLA insert was ligated into the appropriate expression vector pKG4 (obtained from Prof. A. Townsend, Institute of Molecular Medicine, Oxford). Due to a lack of compatible restriction sites, the HLA insert required subcloning into a shuttle vector (pBluescript II KS-, Stratagene). Using EcoR1 and BamH1 restriction sites incorporated into the HLA-A specific primers, the full length gene was cut and subcloned into a similarly digested pBluescript vector. Clones were screened by PCR to establish presence of the HLA insert. Compatible restriction sites (HindIII/BamH1) now existed for subcloning out of the shuttle vector into the pKG4 expression vector.

2.3.2 Maxiprep preparation.

Following transformation of MOSblue competent cells (Amersham Life Sciences), clones were screened for the presence of the HLA insert. A seeding culture of cells containing the positive clone was then used to inoculate 1L of L broth containing 50µg/ml Ampicillin. The cells were grown overnight at 37°C on an orbital shaker. Plasmid DNA was extracted using an alkaline lysis method (Birnboim, et al 1979). Cells were pelleted and supernatant discarded. The pellet was resuspended in 40ml of Buffer-1 prior to alkali lysis by addition of 80ml Buffer-2. This was then neutralised through the addition of 40ml of Buffer-3. Buffers 1-3 are given in Table 2.5. The resultant precipitate was centrifuged down and the supernatant passed through a muslin gauze. 96ml isopropanol was added and DNA precipitated at -20°C for 30 minutes prior to centrifugation. The pellet was then resuspended in 10ml TE to which was added 10g CsCl and 1ml Ethidium Bromide solution. This was then distributed into 4 polycarbonate tubes compatible for use with the Beckman TL100.3 rotor. The plasmid DNA was spun through the CsCl gradient overnight at 80,000rpm at 20°C. The lower band containing plasmid DNA was then carefully removed, and ethidium bromide extracted with amyl alcohol. The plasmid DNA was then dialysed in 2L of TE at 4°C overnight, and the DNA concentration determined by spectrophotometry.

Buffer 1	Buffer 2	Buffer 3
50mM Glucose	0.2M NaOH 1% SDS	3M Na Acetate pH4.8

Table 2.5 Reagents used for maxiprep alkali lysis extraction of plasmid DNA.

2.3.3. Transfection of plasmid DNA into C1R cells.

HLA-A genes were expressed in CIR for the characterisation of peptide motif bound to a particular allelic variant. This was achieved through the transfection of the HLA-A gene through the vehicle of the pKG4 expression vector into the low expression C1R cells. The transfection protocol followed is described below;

20µg of CsCl purified pKG4/HLA was linearised through cutting with 40 units of the restriction enzyme ScaI at 37°C for 4 hours or overnight. The linearised DNA was recovered under sterile conditions by ethanol precipitation with 2 volumes absolute ethanol and 0.3 volumes 3M Na Acetate. This was washed with 70% ethanol, dried and resuspended in 50µl TE.

As an additional control, the pKG4/HLA was also cut with BamH1 and HindIII restriction enzymes. This digestion produced a clean doublet corresponding to the HLA full length insert and the cut expression vector.

Following screening on an agarose gel, to check that the expression vector had indeed been cut cleanly, the 20µg of pKG4/HLA DNA was mixed with 0.8ml of PBS containing 10^7 washed C1R cells. This mixture was added to a cooled electroporation cuvette, and the electroporation performed at 250V, 500µFd. A control was also electroporated in which no DNA was added to the cells. After electroporation, the cuvettes were placed on ice for 10 minutes prior to transfer of the cells to 20ml prewarmed RPMI1640/10%HIFCS. This was distributed into two 75ml tissue culture flasks and incubated at 37°C/5%CO_2 overnight. The following day the electroporated cells were placed in 2mg/ml G418 selection. Of the two control flasks, only one was put in selection to ensure viable cells survived post electroporation. The transfected cells were then incubated for a further 2-3 weeks in until it became obvious that they were growing well in

selection. Still in selection, new flasks were seeded to grow out the live cells. Frozen stocks were made from the fresh flasks and the transfectants checked for expression by FACs analysis.

The volume of tissue culture was further expanded and transferred to roller bottles. At this step, because culture was entering the scale of litres, selection was stopped and the percentage of HIFCS was gradually reduced from 10% to 4%. Following a period of no selection, cells were checked for expression of the transfected HLA-gene. Approximately 25L of cells were cultured yielding in the region of 10^{11} cells. These were washed in PBS and the pellet stored frozen at -70°C.

2.3.4. FACs analysis of transfectants.

Expression of HLA on transfected cells was determined by FACs analysis. In brief;

- 1-2ml (10^6) of viable cells were washed in PBS.
- 100µl of W6/32 mAb (@30µg/ml) was added to the cells and incubated on ice for 30 minutes. A negative control containing no mAb was also prepared.
- Cells were washed twice in cold PBS.
- 75µl of a 1:40 diluted FITC conjugated goat anti mouse were added to the cells and incubated on ice and in the dark for 30 minutes.
- Cells were washed three times in cold PBS.
- Cells were resuspended in 500µl of 1% paraformaldehyde in PBS.
- Cells were stored at 4°C in the dark prior to Facscan analysis.
- Cells were analysed for one colour fluoresence (FL1) with side scatter (SSC).

Controls containing untransfected C1R cells, and B cell line expressing the allele of interest were also stained and analysed.

2.3.5. Serological analysis of transfectants.

Transfectants were further screened for expression by the microlymphocytotoxicity assay (Terasaki, et al 1964; Bodmer, W.F. et al 1979).

In brief, this involved incubating stained viable cells with a panel of HLA specific antisera in Terasaki plates for 30 minutes at room temperature. Following this period, complement was added to the plates and incubation continues for a further hour. The complement requires bound antibody for the cell to be lysed. This was detected by the uptake of a counter red flouresent dye (ethidium bromide, propidium iodide) to enable discrimination between viable and dead cells. The Terasaki plates were read and scored on an inverted UV microscope and the results interpreted accordingly.

2.3.6. Peptide elution and sequencing.

Each transfectant was cultured until approximately 10^{11} cells were available (about a 30ml cell pellet) The cell pellets were washed in phosphate buffered saline and frozen at -70°C. Endogenously processed peptides were then eluted from each of the four transfectants. Peptides were eluted using the trifluoroacetic acid (TFA) method, which was performed essentially as described previously (Falk, et al 1991).

In brief, the cell pellets were lysed in an NP40 based detergent buffer. The cells were then homogenised, ultrasonicated and ultracentrifuged. The supernatant was removed without disturbing the upper lipid phase. The supernatant was filtered prior to being run overnight, through a glycine column and subsequent W6/32 (anti-HLA class I) antibody column. The glycine column acted as a control which determined the proportion of non-specific material bound. The gel in the columns was then resuspended in TFA to elute peptide. The eluate was subsequently ultrafiltrated and peptide fractions separated by reverse phase HPLC (Pharmacia SMART system).

The sequence of the peptide was then determined through Edman degradation sequencing performed with an Applied Biosystems protein sequencer.

2.5. IEF analysis.

IEF analysis was used to discriminate between serologically identical HLA molecules dependent on their comparative charge differences resulting

from their amino acid composition. The basic method used was as described for the 1987 International Histocompatibility Workshop. (Yang, 1989b). A brief description of the protocol is described below;

- Culturing the HLA expressing B cell line in Methionine free medium.
- The cells were labeled with S^{35}-methionine.
- Cells were washed, treated with a protease inhibitor (PMSF) and lysed with an NP40 lysis buffer.
- The lysed cells were then precleared with washed pansorbin.
- The class I specific mAb W6/32 was used to immunoprecipitate the lysed cells.
- The immunoprecipitate was then incubated with Protein A prior to neuraminidase treatment.
- The sample was then suspended in sample buffer and loaded onto into an 30% acrylamide-bis gel containing ampholines (pH3.5-10, pH5-7, pH6-8).
- The gel was run at 10mA for 16 hours.
- The gel was then fixed in 10% acetic acid, washed and treated with 'Amplify'.
- Following drying, the gel was exposed with X-ray film overnight. The film was developed and the IEF banding interpreted.

Aspects of the different methods described above are discussed in the following chapters.

Chapter 3

Establishing the basic method for PCR typing the HLA-A gene.

3.1 Introduction.

This chapter describes the general principles followed in designing the basic PCR-SSP typing system for identifying broad specificities of the HLA-A locus. To achieve this objective, several key areas had to be resolved;

(i) To start, it was necessary to collect all available Class I sequences into one alignment database. This was the fundamental information required to begin designing the PCR-SSP reactions. As new alleles were described, their effect on the specificity of the PCR reactions in the typing panels had to be reviewed.

(ii) Primer pair combinations had to be devised which specifically amplified the required specificities such that no known alleles were missed.

(iii) It was necessary to determine the correct stringency of the PCR to exclude both false positive and negative amplification.

(iv) Conditions had to determined which allowed a panel of PCR reactions, each defining a different specificity, to be performed simultaneously, under a standard set of conditions.

(v) The success of the PCR reaction was confirmed through the incorporation of an internal control PCR reaction.

(vi) The PCR-SSP reactions were validated against DNA obtained from a panel of well characterised or preferably sequenced cell lines.

As described below, establishing the general conditions for the specific amplification of HLA specificities provided the means for design of additional SSP combinations, allowing increasing resolution at the HLA-A locus. Some of the PCR-SSP reactions in the typing panels required alteration over time, in response to new HLA-A alleles found. Parameters within the PCR were also altered, to increase stringency and to accommodate changes in PCR machines.

3.2 Establishing the parameters.

3.2.1 Sequence alignments.

One important prerequisite for designing a PCR based typing system is first to have access to a substantial alignment of the available HLA class I sequences. In designing specific PCR reactions for HLA-A specificities, because of the great amount of shared polymorphism between many class I alleles, not only must other allelic sequences within the locus be considered, but so must HLA-B, -C, and Class I psuedogene sequences (see Chapter 1). This thesis was initiated at a time when the number of published HLA class I sequences (Zemmour and Parham, 1991), was sufficient to begin design of a DNA based HLA-A locus typing system.

Sequence alignments (Zemmour and Parham, 1991; Zemmour and Parham, 1992; Arnett and Parham, 1995) published in conjunction with Nomenclature reports, provided a list of known HLA Class I alleles in a form which allowed polymorphic differences to be readily distinguished. Being able to compare nucleotide sequences in this fashion, allowed sequence specific primers (SSP) to be designed, which when used in combination, allowed specific amplification of a particular sequence.

3.2.2 The use of ARMS; designing SSP combinations.

The basis of PCR-SSP for HLA typing, is specifically to amplify a particular HLA specificity, so distinguishing it from other specificities. There are two major considerations; firstly how to make SSPs specific for the polymorphism they are to detect, and secondly how to combine SSPs in PCR reactions that amplify only the required alleles.

The basis of primer design in this study was the amplification refractory mutation system (ARMS) (Newton, et al 1989). This system inhibits nonspecific PCR amplification when a mismatch exists between template DNA and the 3' residue of one of the primers in the PCR reaction (see figure 1.9), under the correct conditions of stringency. This potentially allows discrimination between two sequences on the basis of a single base pair difference (Newton, et al 1989).

For the definitive amplification of a particular HLA-A specificity, two SSP's were designed, each to identify a point of polymorphism, both of which were found together in the target sequence and not in any other specificities. In other words, the SSP combination had to be unique to the target sequences, with at least one designated polymorphism defined by the SSP's missing in all other sequences. A PCR using this combination of two ARMS designed SSP's under the correct stringency should amplify only the target specificities.

The identification of a particular HLA type in a PCR reaction requires a combination of SSPs that amplify only those alleles containing both polymorphisms along the length of their sequence, as defined by the primers. If a unique polymorphism exists within a sequence, then this provides a good site to which an SSP could be designed. However, such unique sites are not always present, particularly in Class I alleles. The required specificity is then reached through the combination of SSPs, defining two points of polymorphism in a sequence which differentiate it from other HLA types. In designing SSP, it is important to look at other available HLA sequences and pseudogenes, since certain polymorphic combinations may not be confined to a particular locus.

The majority of polymorphisms in HLA-A and other Class I alleles lie within exons 2 and 3 of the gene, and this therefore dictates the positioning of the SSP's. Generally within this protocol, each specific PCR reaction used a primer which identifies a polymorphism in exon 2 in combination with a primer identifying a polymorphism in exon 3. The exon 2 primer extended to become the *cis* strand, while the exon 3 primer extended to become the *trans* strand of the PCR product. Usually, PCR reactions were performed from genomic DNA, and amplification therefore included a 240bp intronic region between the exons 2 and 3. Knowing the distance between priming sites allowed a defined product size to be expected, so allowing greater confidence in interpretation. PCR-SSP reactions were also performed on cDNA with the loss of intronic sequence reducing the size of expected product.

An initial exploratory PCR-SSP reaction attempted amplification of HLA-A*01. Two *cis SSP* in exon 2 were tested in combination with two *trans* SSP in exon 3. Combinations of either AL#1 or AL#2 with AL#A (chapter 2, table 2.2 & 2.3) seemed to be specific for HLA-A1. Combinations with the second *trans SSP* failed, and this was found to be due to a sequence error made in the design of the primer. The components of the PCR were as stated in Chapter 2 including the incorporation of internal control primers in the reaction (chapter 2 table 2.4). The cycling parameters used at this time on a Hybaid Thermal Cycler were;

(95°C-1min, 55°C-1min, 72°C-2min)30 cycles.

Although conditions were not always optimal due to the occasional presence of nonspecific PCR products, the ability specifically to amplify HLA-A*01 provided the ground rules for the design of a basic panel of reactions. The reaction panel covered all the HLA-A locus alleles known at that time, in a generic manner based on broad serological specificities. This initial panel is depicted in Table 3.1, which also shows the perceived specificity of each SSP based on the published HLA-Class I sequences at that time. The observations made with this initial set of reagents provided the basis upon which the general parameters of the method were defined making possible further development.

HLA-A locus Specificity	coding primer	non-coding primer	Specificity of coding primer	Specificity of non-coding primer	band size (bp)
A1	AL#2	AL#A	***A1*** A26 A29	***A1***	510
A2	AL#3	AL#B	***A2***	A9 ***A2*** A6802 A69	413
A3	AL#6	AL#D	A11 ***A3*** A28 A29 A30	***A3***	547
A9	AL#5	AL#B	***A9*** HLA-G	***A9*** A2 A6802 A69	589
A10	AL#4	AL#C	A28 ***A10*** A33	***A10***	436
A11	AL#6	AL#E	***A11*** A3 A28 A29 A30	***A11***	574
A28	AL#4	AL#H	***A28*** A10 A33	A9 A2 ***A28***	466
A19(not A30)	CLI5'	AL#F	HLA CLASS I	***A29 A31 A32 A33***	626
A30	AL#6	AL#G	A11 A3 A28 A29 ***A30***	***A30*** HLA-H	406

Table 3.1 Initial low resolution HLA-A locus typing panel, using a panel of nine PCR-SSP reactions defining eight broad serologically defined specificities (HLA-A30 could not be included in the A19 reaction). The specificity of each SSP is shown with those in bold denoting the specificity common to both primers and so amplified in the PCR reaction. The primer sequences are given in Table 2.2 and Table 2.3.

Internal Control PCR.

To ensure that conditions were complete for PCR to work, an internal control was included in each of the panel reactions. Primers specific for a region of the β-2microglobulin gene (Table 2.4) produced a 330bp product in each of the panel reactions, therefore validating conditions for the PCR. Any reaction which failed to produce either a specific or internal control product, was deemed to have failed. In this protocol, the internal control primers had melting temperatures of around 52°C. This discouraged the internal control reaction out-competing the specific amplification under annealing temperatures of 55°C or above in the PCR. Conversely, any specific HLA amplification would hold advantage and out-compete the internal control amplification under the given conditions. Other methods of PCR-SSP avoided competition from the internal control reaction by lowering the control primer concentration in relation to the specific SSP's (Olerup and Zetterquist 1991; Olerup, et al 1992). The SSP combinations amplified products of a size easily distinguished from the 330bp internal control band.

3.2.3 PCR stringency and conditions for simultaneous amplification.

Initially, each of the primers were designed to be 18 nucleotides in length. Each of the reactions was tested individually with a number of positive and

negative control DNA. Using this initial panel, it became apparent that certain reactions were failing to work under uniform conditions of amplification. This was attributed to the primers not being matched in terms of melting temperature. Certain primers (AL#5, CLI5', AL#C) were therefore extended so that all the SSP had compatible melting temperatures (Thein, et al 1986) of between 56-60°C. In the case of the HLA-A11 reaction, the primer combination of AL#6/AL#E continually failed. This was later attributed to an error in the published sequence. Other combinations were occasionally prone to false positive amplification, and new reactions were designed. Newly published alleles also necessitated the redesign of certain panel reactions to maintain broad specificity, and to ensure no alleles were missed.

The SSP combinations were tested against well characterised control DNA samples (from cell lines with sequenced HLA genes were possible), to validate specificity. Testing each reaction individually against a panel of control DNA identified those reactions prone to cross-reactivity or failure. A number of alternative PCR-SSP combinations were designed for these problem specificities, and alternative reactions were defined.

However, general problems still existed, particularly with the presence of occasional non-specific extra bands, due probably to mispriming within the PCR. The conditions which determined the success of the PCR were multi-factorial. To see if specificity could be improved, alternative conditions within the PCR reaction were investigated. These included;

(i). The use of alternative PCR buffers.

A number of different PCR buffers were tried with HLA-A locus specific PCR-SSP reactions. These included both Potassium Chloride based and Ammonium Sulphate based buffers obtained from different sources, as listed in Table 3.2. The result of using different buffers showed a general loss of amplification with the KCl based buffer while the Biotaq NH_4 buffer tended to slightly increase the amount of nonspecific amplification. However no improvement was noted over the 'in house' 10x buffer.

(A). Perkin Elmer buffer	100mM Tris-HCl pH8.3, 500mM KCl, 15mM $MgCl_2$, 0.01% w/v gelatin.
(B). Biotaq KCl buffer	100mM Tris-HCl pH8.8, 500mM KCl, 15mM $MgCl_2$, 1% Triton X-100.
(C).* Biotaq NH_4 buffer	160mM $(NH_4)_2 SO_4$, 670mM Tris-HCl pH8.8, 0.1% Tween-20.
(D).* In house buffer	166mM $(NH_4)_2 SO_4$, 670mM Tris-HCl pH8.4, 6.7mM EDTA pH8.0, 0.17% BSA.

Table 3.2. 10X PCR buffers tested for use with HLA-A PCR-SSP typing. * $MgCl_2$ not included in these buffers and therefore added separately to final concentration of 2mM in the PCR reaction.

(ii). Magnesium Chloride concentration

The Mg^{2+} concentration in the PCR reaction is important for primer annealing and specificity, strand dissociation, enzyme activity and formation of artifacts. Identifying the correct concentration of Mg^{2+} in the PCR is therefore integral to specific amplification of target sequences. Several titres of Mg^{2+} concentrations were performed (0.0 - 6.0mM) using standard PCR conditions.

It was noted that loss of both specific and internal control amplification was observed at final Mg^{2+} concentrations below 1.5mM, while concentrations above 3mM caused the specific amplification to be reduced in favour of the internal control amplification. In a number of Mg^{2+} titrations performed during the course of this study, a final concentration of 2mM consistently gave the best results in terms of presence of both specific and internal control amplification with reduced (data not shown).

(iii). dNTP , Taq polymerase, and primer concentration.

The final concentration of dNTP's, Taq polymerase, and primer concentrations used in HLA specific PCR reactions has in general remained constant throughout this thesis. All primers were used at a final concentration of 0.2µM, 1 unit of Taq polymerase was used per 50µl PCR reaction. Each dNTP was used at a final concentration of 200µM, with a slight loss of amplification noted at a concentration of 100µM. These

conditions were used standardly in the PCR reactions.

(iv). PCR temperature cycling parameters.

Adjustment of the PCR cycling parameters provided the major means of altering stringency within the PCR reactions. All the SSP were designed to have similar annealing temperatures in an attempt to allow all the typing reactions to be performed under the same amplification conditions. An increase in stringency within the reaction could be achieved by raising the annealing temperature in the PCR, while lowering the annealing temperature would decrease stringency.

Problems were encountered in terms of both false positive and negative amplification at different annealing temperatures. False positives tended to present either as products of the correct size or incorrect size. The correct size product usually indicated a lack of stringency leading to cross-reactive amplification of HLA specificities (i.e. failure of the ARMS principle). Products of incorrect size indicated low stringency or primers annealing to sites different to that intended. If altering stringency had no effect on abolishing false positives, then alternative SSP sites or combinations were considered.

Lack of stringency was associated with false positive amplification. If the situation was general across many of the reactions, an increase in the annealing temperature generally improved specificity. If the problem is limited to one reaction or SSP, then titration of this SSP should be investigated. Also inclusion of a deliberate mismatch in the SSP, near the 3' residue was found to enhance specificity when cross-reactivity was a problem.

The inclusion of step-down PCR enhanced the specificity of the PCR reaction. Step-down PCR involves the alteration of the annealing temperature through the course of PCR cycling. The initial annealing temperature is set high for the first set of PCR cycles before being lowered (or 'stepped down') one or more times in the subsequent cycles. This in effect means that the PCR reaction starts at high stringency which is lowered as the PCR progresses. High stringency improves the accuracy of primer

annealing and so enhances the specificity of the PCR reaction. However, the use of high stringency alone reduces the amount of PCR amplification, making it difficult to detect product. For this reason, the subsequent cycles in the step-down PCR are performed at lower stringency, at which the amplification is more efficient. Creating highly specific PCR product in the first rounds of PCR ensures that when the stringency is lowered, this product will form the basis of further amplification, so reducing the amount of crossreaction and generation of false positives.

False negatives also appeared under conditions that favoured the internal control reaction at the expense of the specific reaction. It was therefore necessary to find the correct window of stringency within the PCR that allowed specific amplification without the presence of false positives or negatives.

As mentioned above, a step-down approach was adopted for enhancing specificity within the panel of PCR-SSP reactions. It also allowed for stringency without the loss of the internal control amplification. Due to the disparity in annealing temperature between the SSP's and the internal control primers, increasing the PCR annealing temperatures caused loss of internal control amplification. Decreasing the annealing temperature caused either an increase of non-specific amplification or false negatives though the internal control reaction competing out the specific reaction. A compromise PCR cycling program was determined, in which the initial PCR cycles were performed at high annealing stringency (65°C) followed by the remainder of PCR cycles performed at a lower annealing stringency (55°C). This step-down PCR enhanced the specificity of the HLA-A SSP combinations while also permitting amplification of the internal control reaction. Additional 'steps' down can be introduced within the PCR program to fine-tune stringency.

Step down PCR;	(95°C-1min, 65°C -1min, 72°C-2min)5 cycles +
	(95°C-1min, 55°C-1min, 72°C-2min)25 cycles

The use of step-down PCR provided the required stringency to control for false positive and negative amplification. The above step-down PCR was as used for much of the early work in the initial panels using the Hybaid thermal cycler. Another PCR cycling program was determined for use with

MJ Research Thermal cyclers in which PCR was performed in 96 or 192 well plates in reduced volumes;

96 well plate step down ;	(96°C-50s, 70°C-50s, 72°C-50s)5 cycles +
	(96°C-50s, 61°C-50s, 72°C-50s)15 cycles +
	(96°C-50s, 55°C-50s, 72°C-1min)10 cycles

Performance of typing reactions in microtitre plates rather than 0.65ml eppendorf tubes, allowed a larger throughput of samples and reduced reaction volumes (from 50µl to 25µl or 13µl). The change in PCR machines and small reaction volumes also reduced cycling times from 4 hours to 2 hours for the 30 cycle PCR's described above.

(v). Preparation of reagents - consistent typing performance.

A key factor in performing PCR-SSP HLA typing is consistency in performance and results. In establishing the basic method, assembly of the PCR reaction was initially performed on an individual basis, to test for a particular SSP combination or other parameter. Results improved as the conditions for PCR became optimised. However occasional variation was noted on repetition of a typing, which could be attributed to a lack of consistency in preparing each PCR typing individually. To abrogate this problem, once the basic PCR conditions were established, large batches of PCR mixes (containing dNTPs, PCR buffer, $MgCl_2$) were made and aliquoted. Preparing such batches of PCR mixes reduced the problem of variation between each PCR amplification. Also the preparation of larger volumes of PCR mix reduced the variation between batches contributing to a more standardised assay. Incorporating such standardisation of reagents allowed for a more straightforward development of PCR protocols for HLA-A typing.

3.2.4 Validation of the PCR-SSP Low Resolution HLA-A typing panel.

Each of the reactions in the initial low resolution typing panel (LRT) in Table 3.1 was tested individually, to determine specificity. Several SSP combinations required alteration or replacement through the influence of newly identified alleles, cross-reactivity with pseudogenes or other class I alleles, or simply because better SSP combinations existed than those first

considered. The continuous reassessment of the typing panel in response to newly identified specificities was a requirement throughout this thesis.

HLA-A locus Specificity	coding primer	non-coding primer	Specificity of coding primer	Specificity of non-coding primer	band size (bp)
A1	AL#2	AL#A	***A1*** A26 A29	***A1***	510
A2	AL#3	AL#H	***A2***	A9 ***A2*** A28	453
A3	AL#6	AL#D	A11 ***A3*** A28 A29 A30	***A3***	547
A9	AL#5	AL#H	***A9*** A3002 HLA-G	***A9*** A2 A28	633
A10	AL#7	AL#C	HLA-A except A202 A205	***A10***	440
A11	AL#6	AL#I	***A11*** A3 A28 A29 A30	A1 ***A11***	518
A28	AL#6	AL#H	A11 A3 ***A28*** A29 A30	A9 A2 ***A28***	445
A19(not A30)	AL#7	AL#F	HLA-A except A202 A205	***A29 A31 A32 A33***	520
A30	AL#10	AL#G	***A30*** A31	***A30*** HLA-H	446

Table 3.3. The first LRT panel following modification from pilot panel listed in Table 3.1

The first panel to be used for PCR-SSP typing at the HLA-A locus was as described in Table 3.3. To validate the performance of the LRT panel, a small number of DNA samples covering most of the specificities identified in the typing panel was tested blind. The results as compared to serology are shown in Table 3.4 below. The results were highly encouraging with agreement in all but one of the samples. The one discordant result was later found to be due to a wrongly labeled sample and not a problem with reaction specificity.

The LRT panel was further used to type genomic DNA extracted from peripheral blood mononuclear cells (PBMC) and EBV transformed B cell lines. Examples of typings using an LRT panel are given in figure 3.1. Additional validation was achieved through typing DNA obtained from serologically well characterised samples and where possible cell lines that had been sequenced for HLA-A locus alleles. Continual use of the panel for typing of samples provided a constant if limited means of validation of the methods performance, since many samples had already been typed serologically. This continual assessment of performance allowed improved SSP combinations to be designed.

CELL LINE	SSP-PCR HLA-A LOCUS TYPING.		SEROLOGICAL HLA-A LOCUS TYPING.	
HRC.1	1	3	1	3
HRC.4	1	2	1	2
HRC.6	1	9	1	9
HRC.7	10	28	26(10)	28
HRC.23	2		2	
HRC.30	2	19	2	29 (19)
HRC.300*	2	9	1	28
HRC.302	2	19	2	29 (19)
HRC.304	2	19	2	31(19)
HRC.305	2	11	2	11
HRC.309	2	3	2	3
HRC.314	2	19	2	31(19)

Table 3.4. A panel of DNA tested blind by the PCR-SSP LRT reaction panel and compared to their serological typing. Concordance between PCR-SSP and serology was noted in all samples except *HRC.300, which was later shown to be a wrongly labeled sample.

The typing panel for low resolution typing listed in Table 3.5 provides a more current list of reactions for this level of resolution. An additional reaction has been added to the nine broad reactions to define HLA-A*80, which is a distinct specificity not identified in the other reaction groups.

The basic panel of the ten reactions listed in Table 3.5 was supplemented by two control reactions which divide the HLA-A locus specificities into two. Amplification with reaction 11 and or 12 can be used to confirm the observation made in the basic panel in a manner similar to using Bw4, Bw6 assignment at B locus. The combined use of SSP's AL#30 and AL#31, specific for HLA-A was necessitated through observations made with the initial choice of coding SSP, AL#7 for these reactions, a cautionary tale with regard to the amplification of pseudogenes. The combination AL#7/AL#L potentially amplifies certain HLA-H alleles, which has resulted in false positive amplifications on several occasions. Interestingly, the presence of these false positives has coincided with HLA-A2 being present in the

haplotype, suggesting potential linkage disequilibrium between HLA-A2 and the relevant HLA-H alleles. The important point to note from this is that not only must other Class I genes be considered in primer design, but also any Class I pseudogenes. An additional general HLA-A locus specific PCR (AL in Table 3.5), amplifying the majority of exons 2 and 3 can be used for additional typing through a nested approach, as discussed in the following chapter.

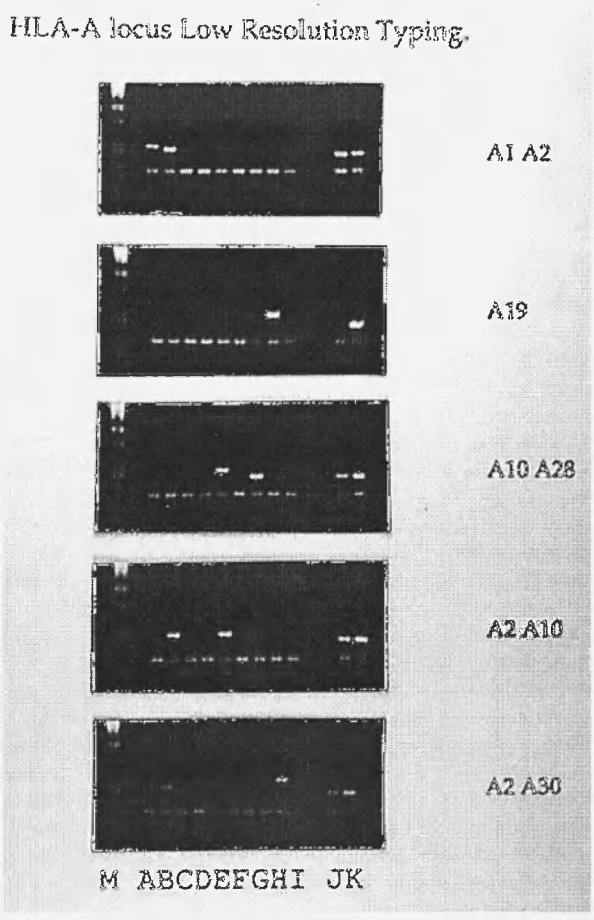


Figure 3.1. Five examples of low resolution typing at HLA-A locus. A panel of nine reactions define as follows;

Lane A; A1 (AL#16/AT)

Lane B; A2 (AL#13/H)

Lane C; A3 (AL#7/D)

Lane D; A9 (AL#8/H)

Lane E; A10 (AL#7/C)

Lane F; A11 (AL#6/I)

Lane G; A28 (AL#6/H)

Lane H; A19(not A30) (AL#7/F)

Lane I; A30 (AL#12/G)

Reaction J and K act as controls with reaction J (AL#30.31/H) amplifying A2,A9,A28 and reaction K (AL#30.31/L) amplifying A1, A3, A10, A11, A19. Lane M is a size marker.

Tube No.	Specificity	Coding Primer	Non-coding Primer	size
1	A1/36	AL#16	AL#AT	527
2	A2	AL#13	AL#H	487
3	A3	AL#7	AL#D	626
4	A9	AL#8	AL#H	470
5	A10	AL#38	AL#AE	649
6	A11	AL#6	AL#I	518
7	A28	AL#6	AL#H	445
8	A19 (not A30)	AL#38	AL#F	580
9	A30	AL#12	AL#G	560
10	A80	AL#54	AL#BK	543
11	A2 A9 A28	AL#30&31	AL#H	446
12	A1/36 A3 A10 A11 A19 A80	AL#30&31	AL#L	446
AL	A locus	CLI5'	AL#AW	780

Table 3.5. A low resolution typing panel in the context of recently identified alleles. Additional control reactions 11 & 12 provide validation of the typing result.

3.3. Discussion.

In this chapter, the development of a fundamental system for analysis of HLA-A polymorphism has been described. The determination of a workable system identified certain key factors which provided points of adjustment to the specificity of the system and also enabled the simultaneous performance of a series of PCR reactions each defining a different HLA-A specificity. Designing primers on the basis of ARMS (Newton et al, 1989) provided amplification specificity under the correct conditions of stringency achieved through the PCR cycling profile, as enhanced through the use of 'step-down' PCR. The ability to simultaneously amplify several different PCR reactions under a standard set of conditions was achieved through assuring all SSP had compatible annealing temperatures. A further important parameter was the 'batch' preparation of reagents to allow for consistency between typing results.

The advantage of using such a system became apparent at even this point. Typing from DNA rescinded the need for viable cells as required in serology,

which in turn facilitated easier storage of samples. There was no reliance on non-renewable stocks of antisera, nor the problems of their cross-reactivity. The PCR-SSP system provided the opportunity for a standardised yet flexible system using standard laboratory equipment and renewable reagents.

Having been successful at a low level of typing resolution and having ascertained the relevant parameters to type reliably by PCR-SSP for HLA-A broad specificities, new SSP combinations were developed which increased the resolution offered by this typing approach. The increase in resolution and the approaches taken to acquire allelic definition are discussed in the following chapter.

Chapter 4

Subtyping and allelic definition.

4.1 Introduction.

The basic parameters for defining broad HLA-A locus specificities had been determined in Chapter 3. However the level of resolution offered by the basic low resolution panel was well below that possible through serological determination. Although the PCR-SSP approach at the HLA-A locus already held some advantages over serology in terms of ease of interpretation, standardised methodology and wide array of source material that could be utilised, the degree of resolution offered clearly needed improvement.

It became apparent through the development of the initial typing panel, that additional PCR-SSP reactions could be readily designed, offering subtyping of the broad serological specificities in the basic LRT typing panel. Because of the shared nature of polymorphism present amongst HLA-A specificities, some of the subtyping reactions could simply be designed using new combinations of SSP's already used or tested for the LRT panel. The first aim was to extend the number of specific reactions to attain a level of resolution comparable with that attainable by good serology at the HLA-A locus. It then remained to design additional SSP combinations and develop approaches which would allow allelic discrimination at the HLA-A locus. This chapter covers the developments made in protocols for achieving increasing resolution at the HLA-A locus to the point of defining the allele.

4.2 Determining the common subtypes of HLA-A9, -10, -28 and -A19.

The described low resolution typing system for the Class I HLA-A locus assigns HLA Class I specificities A*01, A*02, A*03, A9, A10, A*11, A28, A19 (except A*30) A*30 and A*80 from genomic DNA. At this level, serology remains more definitive. Expansion of the initial reaction panel has allowed a higher resolution HLA-A locus typing, which for some alleles gives clearer definition than serology. This section describes this extension to the low resolution typing, by defining the common subtypes of A9, A10, A28, and A19.

The design of additional SSP combinations enables the identification of subtypes of the group specific reactions from the low resolution typing. For example, A9 can be split into A*23 and A*24 and A28 splits into A*6801, A*6802 and A*6901. One of the more challenging groups to assign subtypes for was the A19's. Out of necessity, the method already splits A*30 from the other A19's in the low resolution panel (chapter 3), due to a scarcity of shared polymorphism between A*30 and the other members of the A19 group. However, further SSP combinations were created to distinguish between A*29, A*31, A*32 and A*33. This initial exercise in subtyping at the HLA-A locus also described primer combinations that differentiate between three allelic variants of A*30, (A*3001, A*3002 and A*3002) and two A*29 allelic variants (A*2901 and A*2902).

4.2.1 HLA-A9 Subtyping.

Table 4.1 shows the subtyping reactions which define A*23 and A*24, the serological 'splits' of HLA-A9. An A*23 sample would amplify with reactions (a) and (b), while an A*24 would be positive with reactions (a) and (c).

Reaction	specificity	coding SSP	non-coding SSP	size (bp)
(a).	A9	AL#8	AL#H	470
(b).	A*23	AL#8	AL#Q	555
(c).	A*24	AL#8	AL#R	555

Table 4.1 Subtyping reactions defining A*23 and A*24.

4.2.2 HLA-A10 Subtyping.

HLA-A10 subtyping was performed by a pattern of reactions, as the SSP combination designed to identify A*26 also amplified the rare A*43 allele (Table 4.2). The SSP combination (c) distinguished A*26 and A*43, permitting the differentiation of these alleles. Three SSP combinations were required to define A*25 and A*26, and a further combination confirms HLA-A10. The inclusion of the HLA-A10 SSP combination enabled this panel to identify the specificities A*25, A*26, A*43, and A*34/A*66. Figure 4.1 shows results from A10 subtyping with this panel, in which sample 1 types as an A*25, having positive reactions with (a), (c) and (d), sample 2 types as an A*26, being positive reactions with (b), (c) and (d), and sample 3 types as an A*43, being positive with reactions (b) and (d), but negative for reaction (c). An A*34 or A*66 DNA sample would be positive with reactions (c) and (d), (not shown in figure 4.1).

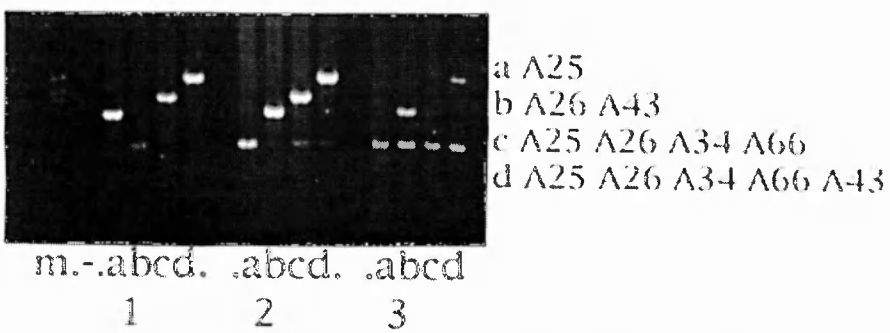


Figure 4.1. Examples of A10 subtyping. Sample 1 types as A*25, sample 2 types as A*26 and sample 3 types as A*43. The constant lower band corresponds to the internal control. SSP combinations are given in Table 4.2

Reaction	specificity	coding SSP	non-coding SSP	size (bp)
(a).	A*25	AL#11	AL#C	398
(b).	A*26 A*43	AL#2	AL#C	397
(c).	A*25 A*26 A*34 A*66	AL#4	AL#C	438
(d).	A10	AL#7	AL#C	497

Table 4.2 Subtyping reactions defining A*25, A*26, A*43, A*34/66.

4.2.3 HLA-A28 Subtyping.

As with the A10 splits, the HLA-A28's were subtyped by reaction pattern. Table 4.3 shows the three SSP combinations, (a) being A*6901 specific, (b) being A*6802 and A*6901 specific while (c) amplifies A*6801, A*6802 and A*6901. A fourth specific combination (d) is also included, to differentiate between A*68 and A69. Reaction (d) also amplifies A*3401 and A*66 and so may not be informative in the presence of these alleles. Although this panel does not include a specific combinations for either A*6801 or A*6802, different patterns of reaction are obtained with the three primer combinations. Thus A*6801 amplifies with primer combination (c) only, A*6802 amplifies with primer combinations (b) and (c), while A*6901 amplifies with all three SSP combinations.

Reaction	specificity	coding SSP	non-coding SSP	size (bp)
(a).	A*6901	AL#6	AL#Y	382
(b).	A*6802 A*6901	AL#6	AL#B	405
(c).	A28	AL#6	AL#H	445
(d).	A*68 A*3401 A*66	AL#6	AL#U	530

Table 4.3 Subtyping reactions defining A*6801, A*6802 and A*6901.

4.2.4 HLA-A19 Subtyping.

Subtyping the HLA-A19 family of alleles has proved to be one of the problem areas for serology. SSP combinations which allow clear distinction between the major HLA-A19 subtypes are shown in Table 4.4. As described previously, HLA-A*30 could not be included in the A19 group specific reaction by PCR-SSP, and requires a separate reaction in the low resolution panel. Therefore, if a low resolution typing panel is performed initially on an A*30 sample, it is not necessary to further subtype for A19 unless the A19 (not A*30) reaction appears positive indicating the presence of another A19 allele.

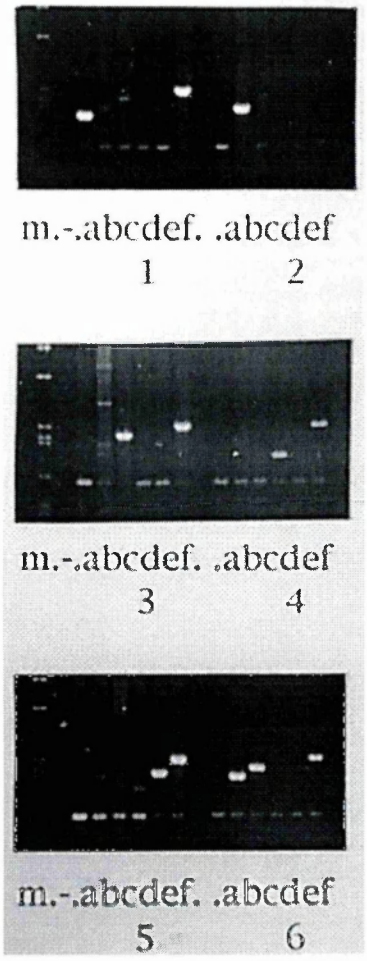


Figure 4.2. Examples of HLA-A19 subtyping to identify specificities; A*29 (sample 1), A*30 (sample 2), A*31 (sample 3), A*32 (sample 4), A*33 (sample 5) and A*30, A*31 (sample 6). The reaction panel consists;

Lane a; A*29
Lane b; A*30
Lane c; A*31
Lane d; A*32
Lane e; A*33
Lane f; A19 (not A*30)

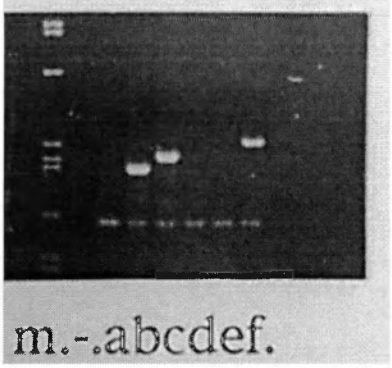
SSP combinations are given in Table 4.4.

Reaction	specificity	coding SSP	non-coding SSP	size (bp)
(a).	A*29	AL#2	AL#F	420
(b).	A*30	AL#12	AL#G	560
(c).	A*31	AL#10	AL#F	480
(d).	A*32	AL#11	AL#F	421
(e).	A*33	AL#4	AL#F	461
(f).	A19(not A*30)	AL#7	AL#F	520

Table 4.4 Subtyping reactions defining A*29, A*30, A*31, A*32, A*33 and A*74 by exclusion.

A19 A30

Figure 4.3a. Low resolution typing of an individual with a serologically problematic combination of A19 specificities, (A*30, A*31). The A*31 in this panel types as an A19 (not A*30). The primer combinations used in this typing panel are listed in chapter 3, Table 3.3.



m.-.abcdef.

Figure 4.3b. A19 subtyping of the individual identifying both A*30 and A*31. The SSP combinations of panel are are listed in Table 4.4.

Figure 4.2 shows six examples of A19 subtypings using a panel of combinations to type for specificities A*29, A*30, A*31, A*32, and A*33, together with the group specific reaction for A19 (except A30). Although not ideal, the A19 split HLA-A*74 can also be determined by amplification with

reaction (f) and exclusion of A*29, A*31, A*32 A*33. A specific PCR reaction is required to identify A*74 with confidence. In figure 4.2, samples 1 to 5 typed as A*29 to A*33 respectively.

Sample 6 typed as an A*30, A*31. Individuals who are A*30, A*31 are extremely difficult to type by serology, due to cross reactivity within the typing sera. This individual was initially typed by serology, through segregation of the two antigens in a family study. Figure 4.3 shows the ease of discrimination of HLA-A*30 and -A*31. Firstly in the low resolution typing panel (top panel in figure 4.3a) in which the A31 is identified as A19 and in subsequent A19 subtyping (bottom panel in figure 4.3b), in which both A30 and A31 are definitively identified. This example demonstrates the virtues of this PCR approach in defining serologically problematic specificities.

4.2.5 HLA-A*30 subtyping.

The typing panel listed in Table 4.5 was primarily designed to discriminate between the common allelic variants of HLA-A*30, A*3001 and A*3002. The results from figure 4.4 show sample 1 typing as an A*3001, being positive with the A*3001 specific primer combinations, (a) and (b), while sample 2 typed as an A*3002, being positive with primer combinations, (c) and (d). Both samples were positive with the additional HLA-A*30 combinations (e) and (f). At the time of design of the initial subtyping panel, only A*3001 and A*3002 had been described, and A*3003 and A*3004 were not considered. The influence of A*3003 and A*3004 on the specificity of the subtyping panel is included in Table 4.5. It would have been possible to determine A*3003 and A*3004 since A*3003 would have presented by amplification with reactions (d) and (f) and A*3004 would have amplified with reactions (e) and (f). Allele specific discrimination of the A*30 group of specificities is described later in this chapter (section 4.4.3).

Figure 4.4. A*30 subtyping by a panel of six reactions as listed in Table 4.5. Sample 1 types as A*3001 and sample 2 types as A*3002.

Reaction	specificity	coding SSP	non-coding SSP	size (bp)
(a).	A*3001	AL#6	AL#K	518
(b).	A*3001	AL#12	AL#K	672
(c).	A*3002	AL#10	AL#T	557
(d).	A*3002, A*3003	AL#12	AL#T	671
(e).	A*3001 A*3002 A*3004	AL*10	AL#G	446
(f).	A*3001 A*3002 A*3003 A*3004	AL#12	AL#G	560

Table 4.5 HLA-A*30 subtyping reactions defining the alleles A*3001, A*3002 , A*3003, A*3004.

4.2.6 HLA-A29 Subtyping.

Two variants of HLA-A*29 have been sequenced so far. The variants differ by only one base pair in the whole of exons 2 and 3. The '2901' SSP has been designed to detect this one base pair difference, which when used in the given combination (Table 4.6), produces a specific amplification for A*2901. The SSP combinations were used to screen known A29 samples in an attempt to locate an A*2901 variant. In figure 4.5, samples 2 and 5, were both positive with the A*2901 specific primer combination, whilst samples 1, 3 and 4 failed to amplify with this primer pair, typing these cell lines as A*2902.

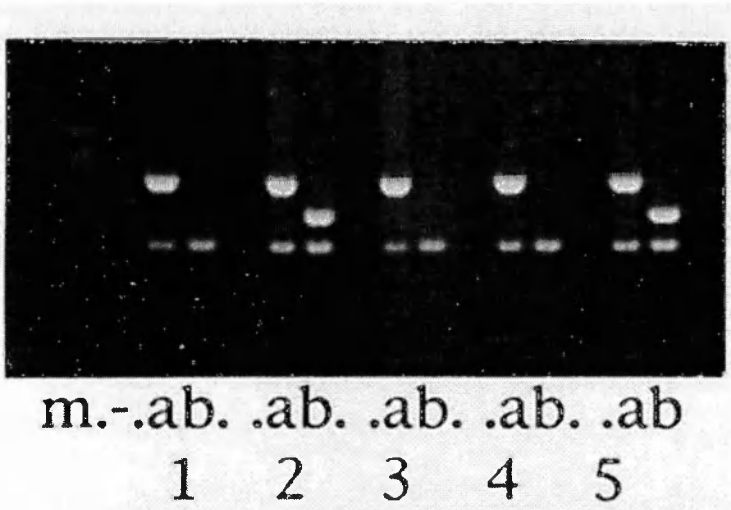


Figure 4.5. Examples of subtyping for A*2901 and A*2902 using SSP combinations listed in Table 4.6. Samples 2 and 5 type as A*2901 and samples 1, 3 and 4 type as A*2902. The lower band present in each typing reaction corresponds to the internal control at 330bp.

Reaction	specificity	coding SSP	non-coding SSP	size (bp)
(a).	A*29	AL#2	AL#F	420
(b).	A*2901	AL#6	2901	367

Table 4.6 HLA-A*29 subtyping reactions defining the alleles A*2901 and A*2902.

The above typings clearly illustrate the ability of ARMS-PCR to differentiate between alleles that differ from each other by a single base pair.

4.2.7 Conclusions to the primary HLA-A subtyping study.

Section 4.2 investigates the possibility of increasing the resolution of the PCR-SSP HLA-A locus typing system from that initially achieved in defining broad specificities. By using the PCR-SSP approach to identify of the common subtypes HLA-A9, -A10, -A28 and -A19, an alternative method to serological definition has been demonstrated at the HLA-A locus.

In designing SSP combinations to define subtypes, there exists a certain level of flexibility as to the choice of primers as determined by the available polymorphism. Hence, a number of different SSP combinations may be used to define a single specificity. The level of resolution is another point of flexibility. The method had at this point already shown definition at a broad generic serological level to one well beyond serology with the allelic definition of A*2901, A*2902, A*3001, A*3002, A*3003. Subtyping has also

demonstrated that PCR-SSP can discriminate specificities on the basis of a single base pair difference as demonstrated on the A*2901 C→G A*2902 substitution found at nucleotide position 33 in exon 3.

Many of the HLA-A locus alleles common in Caucasoids will be detectable by this protocol. However, instances have occured, especially when typing other ethnic groups, when the presence of a novel sequence may cause a discrepancy in the typing result. Discrepancies between serological and PCR typings may point to the existence of a new allelic variant. Further investigation using different primer combinations for the relevant specificity, may offer a prediction of the site of polymorphism in the new sequence with respect to the other variants. The development of this approach in characterising new alleles is discussed elsewhere (chapter 6.2). Alternatively, inclusion of broad generic reactions in addition to the subtyping reactions may indicate discrepancies (e.g. a positive amplification with the generic reaction but all the subtyping reactions type negative). Such examples emphasize the need to be constantly aware of new sequence data to ensure that the typing panel has valid specificity and covers all HLA-A locus alleles. The use of the additional SSP combinations AL#30.31/AL#H and AL#30.31/AL#L contribute a partial level of control in such situations.

The ability of the PCR-SSP approach to distinguish HLA-A locus alleles in different heterozygous combination is a significant attribute of this tissue typing system. However, using the SSP combinations described in this section, there remain some instances in which a definitive typing would not be possible in defining certain heterozygous combinations. In subtyping for A28, A*6801 could not be distinguished from A*6802 if present together in a heterozygote. The described A10 subtyping panel would not be able produce a definitive typing if combinations of A*26 and A*43, or A*43 and A*34/A*66 were present together in a heterozygote. Also, in splitting A*29, the given SSP combinations would not distinguish A*2901 if present together with A*2902 in the same individual. Although such antigen combinations will be rare in the general Caucasoid population, it remains important to be aware of their existance. The design of SSP combinations which offer a more concise typing without ambiguity provided the next area for improvement of the PCR-SSP method.

This section has concentrated on defining the main variants or 'splits' of the broad specificities identified by the low resolution typing panel with emphasis placed on specificities present within the Caucasoid population. However, it is also important to be aware of alleles found in other ethnic groups. A subsequent objective was the extension of the method to cover alleles found in black and oriental populations.

4.3. Validation of an extended HLA-A locus PCR-SSP panel through typing DNA from the 4AOH cell panel.

To further verify the fidelity of the HLA-A locus PCR-SSP typing system, the method was applied to DNA samples in the Fourth Asia-Oceania Histocompatibility (4AOH) Workshop. The panel of SSP combinations used provided an equivalent, and in some instances superior level of resolution to that attainable to serology. The study was carried out blind under Workshop conditions and the results confirm the method as an accurate means of determining HLA-A locus tissue types.

The 4th Asia Oceania Histocompatibility Workshop (4AOH) typing component offered the opportunity of DNA typing the HLA-A locus of workshop cell lines in a blind study. This study represents the first comparison of DNA based and serological typing methods for Class I alleles to be carried out under the controlled conditions of an HLA workshop. Cells grown from frozen stocks were used to extract genomic DNA using an Applied Biosystems 340A nucleic acid extractor and standard protocols. 48 samples were typed by an initial low resolution panel of 9 reactions, (A1/36, A2, A3, A9, A10, A11, A28, A19(not A30) and A30) with subsequent subtyping of A9, A10, A28 and A19; the method being as described previously in both this chapter and chapter 3. The SSP combinations used were capable of determining specificities as listed in Table 4.7. Following PCR, a sample of each reaction was run out on an agarose gel and the typing result determined by the presence or absence of the appropriate sized product.

LOW RESOLUTION TYPING SPECIFICITIES	SUBTYPING PERFORMED
A1/A36	-
A2	-
A3	-
A9	A*23 A*24
A10	A*25 A*26 A*43 A*34/A*66 A*3401
A11	-
A28	A*6801 A*6802 A*6901
A19(not A30)	A*2901 A*2902 A*31 A*32 A*33
A30	A*3001 A*3002 A*3003

Table 4.7. List of HLA-A locus specificities tested for in both the Low Resolution Panel and subsequent subtyping panels.

Examples of low resolution HLA-A locus typings are given in figure 4.6. Sample 003 typed as A*01, sample 011 typed as an A*02, A19(not A*30), sample 012 typed as an A*02, A*11, and sample 026 typed as A*03, A9. Each reaction also included a positive internal control primer set, amplifying a 330bp region of the β-2 microglobulin gene. This product can be seen as the lower band present in each of the panel of nine reactions of the four typings shown in figure 4.6.

Results of the HLA-A locus typings in the 48 workshop samples (containing subtypings) are given in Table 4.8, which also compares data obtained by the PCR-SSP method with the Workshop "Best characterisation of 4AOH cell

panel" (Degli-Esposti, et al 1993). This allowed the determination of the best allele assignment as typed by the most discriminating technique for that specificity and includes typing designations reached by this PCR-SSP method. Agreement was reached between PCR-SSP typings and the Workshop results in 43 of the samples. The remaining 5 samples produced results that did not concur with the workshop designation. Sample 013 was typed by PCR-SSP as an A*02, A*11, but only as an A*02 in the workshop results. Following the release of the Workshop results, the cell was typed retrospectively by serology and the typing suggested the presence of a weak A*11, but the results were not conclusive. Sample 021 was typed from DNA as an A*11 A*6801 as compared to the Workshop result of A*02, A*11. The sample was retrospectively typed by serology, confirming the PCR-SSP A*6801 designation.

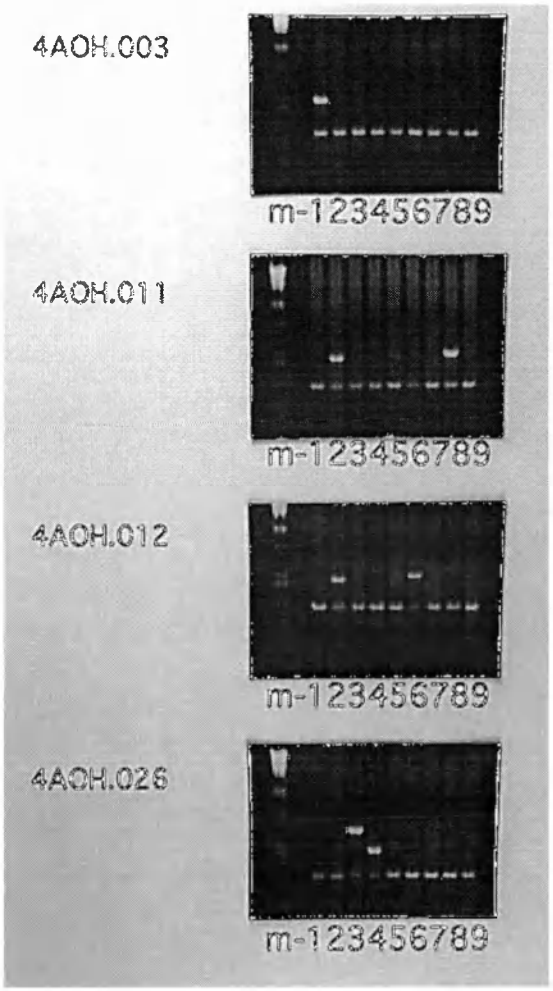


Figure 4.6. Four examples of HLA-A locus low resolution typing of DNA samples from the 4AOH workshop panel. The typing panel consists of;

Lane 1; A1
Lane 2; A2
Lane 3; A3
Lane 4; A9
Lane 5; A10
Lane 6; A11
Lane 7; A28
Lane 8; A19 (not A30)
Lane 9; A30

The panel is as described in figure 3.1 in chapter 3.

4AOH.003 types as A1, 4AOH.011 types as A2, A19(not A30), 4AOH.012 types as A2, A11 and 4AOH.026 types as A3, A9. Additional subtyping was performed to identify the specificities listed in Table 4.8.

4AOH SAMPLE No.	SSP TYPINGS A	A	4AOH TYPINGS A	A
001	2		2	
002	19 (33)		19 (33.1)	
003	1		1	
004	2		2	
005	11	19 (33)	11	19 (33)
006	3	9 (24)	3	9 (24)
008	2	3	2	3
009	10 (26)		10 (26.1)	
010	2	19 (2902)	2	19 (2902)
011	2	19 (31)	2	19 (31)
012	2	11	2.2	11.2
013	2	11	2	
014	2	19 (31)	2	19 (31)
015	2	9 (24)	2	9 (24)
017	2	19 (31)	2	19 (31)
018	9 (24)		9 (24)	
019	1		1	
020	2		2.2	
021	11	28 (6801)	2	11
022	1		1	
024	2		2.4	
026	3	9 (24)	3	9 (24)
027	9 (24)		9 (24)	
028	2	10 (3401)	2	34
029	2		2	
030	19 (31)		19 (31)	
031	2		2	
033	2		2	
037	3		3	
038	3		3	
039	1		1	
040	3		3	
045	1		1	
052	11		11.2	
059	2		2.2	
068	19 (31)		19 (31)	
070	2		2.2	3
072	2		2.2	
073	2		2	
077	2		2.2	3.2
141	1		1	
200	2		2.4	
204	2		2	
208	1		1	
211	3002		30	
016*	11	19 (2902, 33)	11	
032*	1, 2	9 (24)	9 (24.1)	
036*	3, 9 (24)	10 (26), 30	3.2	9 (24.1)

Table 4.8. Comparison between ARMS/SSP DNA typing and 4AOH Best Characterisation typing results for HLA-A locus. * Indicates DNA samples with contamination.

In the remaining three samples (016, 032, 036), more than two HLA-A locus alleles were identified in each sample. These typing differences were shown to be a result of DNA contamination and not an intrinsic problem of the PCR. PCR-SSP typings of these samples consistently identified additional HLA-A locus alleles. Sample 016 was typed in the Workshop as A*11. The results revealed the additional presence of A*2902 and A33. Serology could only confirm the additional presence of A*29 in this cell line. Sample 032 was reported in the Workshop as an A24.1. Typing by PCR-SSP found a further two alleles (A*01 A*02). Subsequent serology found A*01 A*02 but failed to see A24. Following disclosure of the results it was noted that Sample 032 corresponded to the well defined cell line DKB (A24). Typing a separate source of this DNA from our own stocks confirmed the Workshop result, indicating that the sample typed initially in the study was contaminated. Sample 036 (known better as L0081785), typed as A*26 A*30 in addition to the reported typing of A*03 A24.1. Serology confirmed the workshop result. Repeating the typing with DNA obtained from an independent source, also confirmed the workshop result, whilst the anomalous DNA sample reproducibly gave the additional specificities. Similar instances of contamination have been previously observed, and this highlights the need for extreme care in the preparation of DNA samples.

In this blind study, the degree of concurrence between the ARMS PCR-SSP typings and Workshop results confirms that HLA Class I DNA typing is both feasible and reliable. Even in the instances where disagreement existed, this could be explained by sample contamination or serological mistyping. The IEF data included in the Workshop results (Degli-Esposti, et al 1993), identified some specificities additional to our own, namely A2.2, A2.4, A3.2 A24.1, A26.1, A33.1 in the panel tested. The ARMS PCR-SSP method, however, is capable of typing at an allelic level as shown in identifying A*2902, A*3002 A*3401 and A*6801 in the DNA's tested, so producing far more definitive typings than serology. The full potential of the technique in producing Class I allelic typing was not realised in this study. Further development was therefore necessary to extend this level of allelic typing to all Class I specificities.

As with other DNA based typing methods, ARMS PCR-SSP can only assign allele specificity based on known DNA sequences. The use of group specific

reactions in the low resolution panel minimises the chance that 'new' alleles will be missed by the technique. But new sequences need to be examined to ensure they are detectable correctly by the current panel of SSP combinations, and if not, new SSP combinations have to be designed. The flexibility of the method usually means that more than one SSP combination exists for each specificity or broad group specific reaction, enabling a replacement combination to be found.

In conclusion, the potential of the ARMS PCR technique as demonstrated under Workshop conditions, clearly showed it to be an accurate means of determining the HLA-A locus tissue type from DNA. The successful typing of 4AOH DNA provided the confidence for further development of the method in search of increasing resolution at HLA-A and also the application of a similar approach to HLA-B (Sadler, et al 1994). The use of PCR-SSP for HLA-A locus typing provided future Workshops and studies an additional technique for Class I DNA typing, with a potential well beyond that of serology.

4.4 Nested PCR-SSP for allelic HLA-A typing.

4.4.1 Introduction.

The nature of polymorphism found amongst HLA alleles occurs mainly as a result of different combinations of sequence motifs shared with other HLA alleles at hypervariable regions within the gene. To type at a level of allelic definition is difficult due to a relative absence of unique sequence polymorphisms which would allows for their specific identification in a PCR-SSP system. The lack of unique polymorphism and the prevalence of shared sequence polymorphism amongst HLA Class I genes, increases the complexity for specifically amplifying and identifying a particular allele. One approach to increase the polymorphism available for allelic definition of HLA Class I alleles is the use of nested PCR. The use of nested PCR for HLA typing was first described in relation to class II DNA typing, through definition of DRB1 specificities (Bein, et al 1992).

The general principal of nested PCR-SSP in the context of this study, is the primary amplification of the desired group of specificities using primer sites which flank the polymorphic regions required to distinguish the alleles of interest. This specific first round PCR is then used as template for subsequent PCR-SSP reactions using primer sites found within the initial amplification. The use of a specific first round PCR excludes the interference of other class I alleles which would compromise the integrity of the subtyping reactions. The 'filtering out' of unwanted HLA alleles increases the available polymorphisms from which SSP combinations can be designed which allow differentiation between alleles of highly similar sequence.

The use of nested PCR is also useful in situations where DNA quantity is limiting. As described in Table 3.5 (Chapter 3), the SSP combination CLI5'/AL#AW specifically amplifies the majority of exons 2 and 3 of all HLA-A locus alleles. This PCR product can then be used with the majority of specific SSP combinations used for typing at HLA-A.

The use of nested PCR for subtyping in this thesis, was initially established for defining the allelic variants of HLA-A*03. Further application of the nested approach was used for allelic definition of HLA-A*30 and HLA-A*02. Where required, nested typing was used to define specificities difficult to identify through a one step PCR-SSP approach. This section will describe the development of nested approaches for the discrimination of HLA-A specificities to the allelic level.

4.4.2 Establishing the parameters for nested PCR.

The general parameters for the use of nested PCR in application to PCR-SSP typing at the HLA-A locus was established in relation to A*03 subtyping. Two allelic variants of HLA-A*03 had been sequenced (Bodmer, J.G. et al 1995). As an exercise in screening for A*03 subtypes, it was decided to attempt definition of A*0301 and A*0302.

Definition of these alleles through a one step PCR-SSP proved difficult, since their differentiation relies on identifying polymorphisms found in other class I alleles. The main polymorphisms that discriminate between the two A*03 alleles are found in exon 3 at nucleotide positions 195/6. A*0301 has

nucleotides 'TT' at this site while A*0302 has 'CA'. To define either A*03 variant, it was necessary to combine the discriminatory SSP's with a fairly specific A*03 SSP. The only polymorphic site which has this specificity is as defined by primer AL#D, located some 18 nucleotides 3' of the 195/6 polymorphism in exon 3 and too close to be of use in a one-steps PCR system.

To remove the influence of the potential cross-reactivity from other class I alleles, a nested approach was considered. This would use the A*03 specific SSP combination AL#7/AL#D which flanks the sites of polymorphism which distinguish A*0301 and A*0302. The parameters for PCR-SSP using PCR product as template needed to be determined in the context of the HLA-A locus typing approach established so far.

An approach was investigated whereby the first round A*03 amplification was performed through the standard 30 cycle step-down PCR. A titration of this first round product comprising tenfold dilutions ($1:10^3$ to $1:10^7$ final dilution in the PCR), were then tested in a further 30 cycle step down PCR using nested SSP combination AL#6/AL#Q, specific for A*0301. Using the nested SSP combination, specific amplifications were present at all the dilutions. An A*03 negative DNA was also tested. No specific product was seen and only the internal control was amplified at all the dilutions.

Reaction	specificity	coding SSP	non-coding SSP	size (bp)
(a)	A*0301	AL#6	AL#Q	530
(b)	A*0301	AL#6	AL#AA	519
(c)	A*0302	AL#6	AL#R	530

Table 4.9 Screening panel for identifying A*03 subtyping performed nested from AL#7/AL#D.

A*03 subtyping was performed on a 1:500,000 final dilution of first round product obtained from thirteen A*03 DNA samples, through a 30 cycle step-down PCR using the typing panel given in Table 4.9. Two A*03 negative DNA samples were also included as controls. All thirteen A*03 samples typed as A*0301, and no A*0302 DNA was identified. Neither non A*03

samples gave a positive amplification (data not shown). Unfortunately, sequenced controls were not available for both A*03 subtypes. However, later in this chapter (section 4.5), an A*0302 sample was identified using a similar subtyping panel.

The extent of dilution of the first round product was not really practical for this system. To circumvent this problem, the first round PCR was used at a lesser dilution and the number of PCR cycles reduced. A $1:10^3$ and $1:10^4$ final dilution of A*03 first round product in the PCR were tested with controls against a 20 and 25 cycle A*0301 specific nested step-down PCR using SSP combination AL#6/AL#Q. All of the combinations of PCR cycle number and dilution factor produced clean specific amplifcation. Initially, based on this result, it was decided that a $1:10^3$ final dilution of first round product in combination with 20 cycle PCR would be appropriate for this system. However, in extending this approach to A10 and A19 subtyping, occasional non-specific amplification was observed (data not shown). To counter this problem, the number of cycles within the PCR was reduced to 15. The final conditions used for nested PCR were therefore a final $1:10^3$ dilution of first round PCR product in a 15 cycle amplification. The cycling parameters as used on the Hybaid thermal cycler were;

(95°C-1min, 65°C-1min, 72°C-2min)5 cycles + (95°C-1min, 55°C-1min, 72°C-2min)10 cycles.

Another approach to reduce non-specific amplification was to increase the dilution factor of the first round product. Variation in the dilution factor has proved useful when the first round amplification has been above or below that expected. For example, in the event of a weak first round amplification, a smaller dilution can be used in the subtyping reactions.

The following sections describe the use of nested subtyping with particular analysis of HLA-A*02 and A*30. Finally, the last section describes a comprehensive system for discrimination of HLA-A locus alleles, summarising the previous developments made in this chapter.

4.4.3. Nested PCR-SSP to identify the allelic subtypes of HLA-A*30.

This section describes a nested PCR method to distinguish the allelic variants of HLA-A*30 alleles. This approach was adopted to facilitate easier identification of these alleles without the interference of cross-reactivity from other HLA class I alleles. A first round amplification was used to generate an HLA-A*30/A*3302 specific ARMS/PCR product. At the time this reaction was designed, A*3302 had not been described. The co-amplification of A*3302 would however not interfere in the specificity of the subtyping reactions. This is then used as template for a second round of amplifications in which the different A*30 alleles can each be specifically identified.

Although the A*30 subtyping SSP combinations described earlier in this chapter (Table 4.5), could discriminate between A*30 variants, this was achieved through interpretation of reaction patterns. Typing through reaction pattern may fail to identify certain heterozygous combinations, as indeed is seen in Table 4.5, where an A*3002/A*3003 heterozygote would produce the same reaction pattern as an A*3002 homozygote. To address this situation, it was decided that where possible, specific SSP combinations for each allele or specificity would be incorporated into typing panels, which may require the application of nested PCR.

The initial nested protocol was devised to identify the three variants of HLA-A*30 known at that time. DNA from A*30 sequenced B cell lines were available for each of these three variants and used for verification of the subtyping panel. These included LBF, RSH (A*3001), EJ32B, CR-B (A*3002), and JS (A*3003), with additional local DNA samples complementing the controls. All DNA's were used at 500ng/50µl for the first round PCR reaction.

Reagents and primer concentrations were as described in chapter 2 (Section 2.1.2). The SSP combination used (Table 4.10, SSP sequences in Tables 2.2. & 2.3) for the first round HLA-A*30 A*3302 specific amplification was initially AL#12/AL#Q. This was later modified to AL#12/AL#W to accommodate and characterise a variant A*30 allele, A*3004. These primer pairs are different to the combinations used for identification of A*30 (AL#12/AL#G)

in the low resolution panel (see chapter 3), which does not amplify a sufficient region of exons 2 and 3 for subsequent nested subtyping. The possible co-amplification of A*3302 with SSP combinations AL#12/AL#Q and AL#12/AL#W should not interfere with the A*30 nested subtyping panel of reactions, as all the nested PCR reactions used in the subtyping panel were specific for A*30 alleles. An A*3302 cell line was not available, and so it was not possible to verify its reactivity with the above SSP combinations. The first round amplification consisted of (95°C-1min, 65°C-1min, 72°C-2min) for 5 cycles followed by (95°C-1min, 55°C-1min, 72°C-2min) for 25 cycles. Following the first round amplification, 10µl of PCR product was checked on a gel to identify the presence of a PCR product and confirm the sample as HLA-A*30. The PCR product was diluted 1:100 in dH_2O. 5µl of the dilution was then used in each of the 50µl second round amplifications with the appropriate SSP combinations (Table 4.10). The second round PCR was run on the shortened program consisting of (95°C-1min, 65°C-1min, 72°C-2min) for 5 cycles followed by (95°C-1min, 55°C-1min, 72°C-2min) for 10 cycles. The final PCR product was visualised on an ethidium bromide stained agarose gel, and the result interpreted through the presence or absence of the appropriately sized PCR products.

As part of these studies, a serologically defined A*30 sample was tested and failed to amplify with the initial first round PCR-SSP combination (AL#12/AL#Q). This discrepancy between the serological definition and the DNA suggested the presence of a novel variant of HLA-A*30. Failure to amplify with A*30 PCR-SSP combinations utilising primer AL#Q, provided insight that a polymorphism existed at this site. Further characterisation of this sample by 'PCR gene-mapping' and DNA sequencing (discussed in Chapter 6), defined this A*30 as a new variant and it was subsequently cloned, sequenced and named A*3004 by the WHO nomenclature committee. As mentioned in the Table 4.10, PCR-SSP combinations exist for specific amplification of A*3004, and one of these was included in subsequent panels (see Table 4.13).

Reaction		specificity	coding SSP	non-coding SSP	size (bp)
1st round A*30 PCR		A*3001, A*3002 A*3003, A*3004 A*3302	AL#12	AL#W	719
		A*3001, A*3002 A*3003, A*3302	AL#12	AL#Q	687
2nd round subtyping PCR	(1)	A*3001	AL#6	AL#K	518
	(2)	A*3002	AL#10	AL#T	557
	(3)	A*3001, A*3002 A*3004	AL#10	AL#O	444
	(4)	A*3002, A*3003	AL#12	AL#T	671
	(5)	A*3003	AL#36	AL#T	557

Table 4.10. SSP combinations for HLA-A*30 subtyping by nested PCR. A first round amplification (AL#12/AL#W) is used as template for a second round of subtyping reactions. AL#12/AL#Q was found to not amplify an A*30 variant (A*3004). A specific reaction for the A*3004 can be obtained using any combination of AL#10 with AL#AX, AL#AY, AL#AO, AL#AV or AL#AU as shown in the 'PCR gene mapping' characterisation (Figure 6.1 Table 6.1)

Examples of typing the four variants of HLA-A*30 by nested PCR-SSP, is given in Figure 4.7. Hence, with the panel, A*3001 is identified by amplification with reactions 1 and 3, A*3002 with reactions 2, 3 and 4, A*3003 amplifies with reactions 4 and 5 and the new A*30 variant (A*3004) identified by reaction 3 and failure to amplify with the original first round amplification, AL#12/AL#Q. All four variants can be distinguished and the addition of an A*3004 specific nested PCR-SSP combination to the panel, removes any possible ambiguity in definition of A*30 heterozygous combinations. The four A*30 variants can be defined by four allele specific SSP combinations, however the additional reaction can prove useful for subtype verification.

The identification of new HLA class I alleles provides a constant reason for vigilance in terms of how new sequences may affect the specificity of existing ARMS/PCR reactions. For typing A*30 alleles, we are now aware of two recently identified alleles, A*0102 (Browning, et al 1995) and A*3302 (Kato, et al 1993), for which SSP AL#12 (previously thought to be A*30 specific) also has specificity. Of these two alleles, A*3302 can potentially be co-amplified by

the first round primer combination of AL#12/AL#W. In this study, the subtyping reactions exclude the presence of A*3302, since they identify motifs not seen in this allele. Co-amplification of A*3302 would therefore not affect A*30 subtyping in the methods described. It would however be potentially be misleading to use the SSP combination AL#12/AL#W as A*30 specific, in the low resolution HLA-A locus typing panel as it is in fact A*30 plus A*3302 specific. For this reason we continue to use primer combination AL#12/AL#G, which specifically amplifies all four HLA-A*30 variants, as the generic A*30 SSP combination in the low resolution typing panel (Chapter 3, Table 3.5).

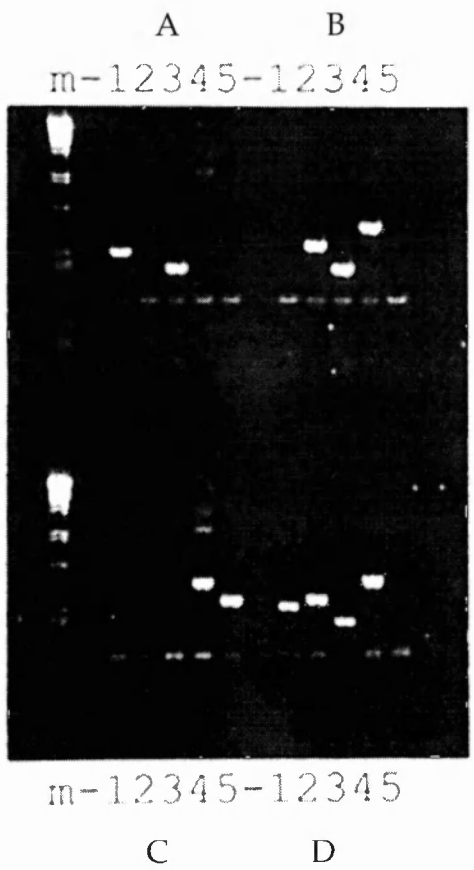


Figure 4.7. Four examples of A*30 subtyping by nested PCR. The panel of 5 subtyping reactions, listed in Table 4.10, are performed nested from 1st round PCR AL#12/AL#W.

Sample A types as an A*3001 being positive with reactions 1 and 3. Sample B types as an A*3002, being positive with reactions 2, 3, 4. Sample C types as an A*3003 being positive with reactions 4 and 5. Sample D types as an A*3001, A*3002 being positive with reactions 1,2,3 and 4. An A*3004 sample would be amplified by reaction 3 alone (not shown).

4.4.4. Nested PCR-SSP to identify the allelic subtypes of HLA-A*02.

HLA-A*02 is present at high frequency in most populations. Seventeen allelic HLA-A*02 variants have been identified at the time of writing. The great majority of genetic polymorphism which defines the allelic variants is

found in exons 2 and 3 of the A*02 genes. These exons encode the α-1 and α-2 domains of the HLA Class I molecules, and variation within the genes may influence the peptide binding specificity of the gene products of each allele. This section describes a method for determining the different HLA-A*02 alleles by use of a nested PCR-SSP approach. This method of HLA-A*02 subtyping permits a more detailed and discriminating study of allelic variation between different population groups as discussed in Chapter 5.

This inability of serology to define adequately functionally distinct allelic variants render it redundant for typing to the T cell level of discrimination Serology can at best identify three A2 specificities (Marsh, et al 1992), and five 1D-IEF (Guttridge, et al 1992). But through DNA sequencing, Sixteen of the seventeen then known A*02 alleles, encode molecules which have different amino-acid sequences and potentially different profiles of immunological function (reviewed in chapter 8). A*0215N is not expressed and is truncated prematurely in exon 4 through a point mutation encoding a stop codon. The relevance of the polymorphic differences that exist between the HLA allelic variants makes their identification desirable. This is particularly true for HLA-A*02, which is found at high frequency in nearly all populations studied.

The HLA-A*02 alleles vary in the main part in exons two and three. Of the 16 expressed HLA-A*02 alleles, only A*0201 and A*0209 are identical in exons 2 and 3, differing only in exon 4 by a single amino acid. All the A*02 alleles show a high level of sequence similarity, containing motifs shared with other A*02 variants as well as other A locus and Class I alleles. The high degree of sequence homology between the HLA-A*02 alleles and the relative lack of unique polymorphisms make identification of the individual alleles potentially difficult. Figure 8.2 (chapter 8) describes the polymorphic differences between the A*02 allelic variants.

A panel of subtyping reactions using SSP designed on the ARMS principle was used to achieve the high level discrimination required to distinguish between the A*02 alleles. These subtyping reactions were run nested from a initial flanking ARMS-PCR which specifically amplifies essentially all of exons 2 and 3 of all A*02 alleles as demonstrated in figure 4.8. The use of an A*02 specific first round reaction excludes the interference of other Class I

alleles in the subtyping reactions. This addition of specificity through a first round reaction enables high resolution typing and discrimination between alleles of highly similar sequence. The system was validated using DNA from sequenced cell lines, and shown capable of distinguishing between all known A*02 alleles.

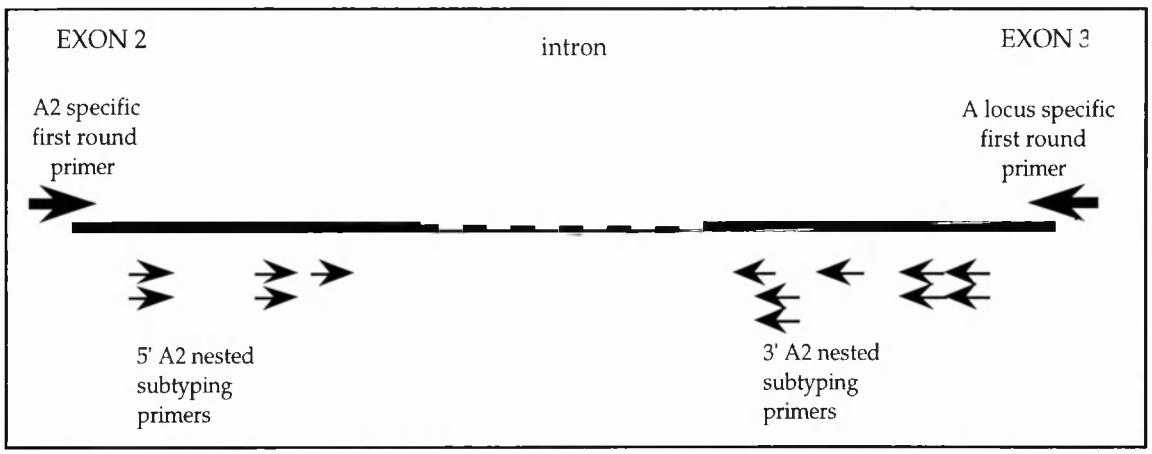


Figure 4.8. Representation of the nested PCR-SSP approach to subtyping the HLA-A*02 specificity. An initial flanking reaction comprising an A*02 specific primer at the 5' end of exon 2 is combined with an A locus specific primer at the 3' end of exon 3. This PCR product which contains many of the polymorphic differences between the A*02 alleles is then used as template for a further round of PCR reactions using primer sites within the initial flanking reaction which determine the A*02 subtype.

Control DNA's representing many of the known A*02 alleles, were obtained from B cell lymphoblastoid lines including JY (A*0201), M7 (A*0202), DK1 (A*0203), RML (A*0204), WT49 (A*0205), CLA (A*0206), KNE (A*0207), KLO (A*0208), OZB (A*0209), XLI.ND (A*0210), KIM.ND (A*0211), KRC033 (A*0212), SLUGEO (A*0213), TUBO (A*0216) and AMALA (A*0217). A*0214 (Krausa, et al 1995) genomic DNA was provided from peripheral blood mononuclear cells (PBMC), from the individual in whom the allele was identified. No DNA was available for the non-expressed A*02 variant, A*0215N.

The identification of HLA-A*02 alleles was achieved in two steps. Firstly a generic 813bp HLA-A*02 specific flanking reaction, using SSP combination AL#37 and AL#AW, amplified all of exons 2 and 3 together with the intervening intron (figure 4.8). Secondly, a panel of SSP combinations lying nested within the first round amplification were used against the A*02

specific PCR product. Initially, only 14 A*02 variants had been identified, which were discriminated by the panel of 13 subtyping reactions given in Table 4.11. SSP sequences are given in Chapter 2, Table 2.2 and 2.3. A*0201 and A*0209, together with A*0207 and A*0215N were not distinguished by nested PCR since the defining polymorphisms are located in exon 4. SSP AL#BJ, used for identification of A*0207, has a melting temperature of 72°C as compared to the standard 58-60°C used in these protocols. The initial SSP used to identify this allele consistently gave weak false positive bands. We overcame this problem by extending the primer to a 21mer, and incorporating a deliberate mismatch (Newton, et al 1989) four nucleotides from the 3' end. This SSP was still compatible with the conditions used for the other SSP's.

A*0207, A*0208 and A*0209 were initially only described as amino acid sequences (Lopez de Castro, 1989). To design SSP's which could be used to identify these alleles, the amino acid sequence was conservatively translated into nucleotide sequence, within the context of other known HLA-A locus alleles. The validity of these SSP sites was then tested in the appropriate PCR combinations against DNA obtained from the sequenced cell lines KNE (A*0207), KLO (A*0208) and OZB (A*0209). In the absence of nucleotide sequence, others have also successfully taken this approach to identify these three alleles by oligonucleotide probing (Fernandez-Viña, et al 1992). The SSP's designed on this basis together with the other SSP's used in the subtyping panel, provided the required specificity in determining the A*02 subtype for these three alleles.

In the initial A*02 subtyping panel described in Table 4.11, identification of the A*02 alleles was achieved by SSP combinations specific for a particular allele or pattern of reactivity. A*0202, A*0203, A*0204, A*0207, A*0208, A*0209, A*0210, and A*0211, were identified by specific PCR reactions. Identification of A*0201, A*0205, A*0206, A*0212 A*0213 and A*0214 relied on interpretation of reaction patterns (e.g. A*0205 was identified by positive reactivity with reactions 'e' and 'f' in Table 4.11). A*0201 and A*0209 were distinguished by an additional reaction from genomic DNA, using an SSP combination amplifying from exon 3 (AL#29) to an A*0209 specific site in exon 4 (AL#AL).

In situations where A*0204 or A*0207 or A*0211 were identified, and no other A locus allele was observed, the additional presence of A*0201 in an A*02 heterozygous combination, could not be discounted by the panel described above. To verify whether A*0201 is present in such situations, additional reactions were required as listed in Table 4.12. Similarly, reactions that detected A*0206 in an A*0206/A*0210 heterozygous sample and A*0212 in an A*0212/A*0213 heterozygous sample were also required. The reactions listed in Table 4.12 were validated against a panel of sequenced A*02 DNA's. Because A*02 is present at high frequencies within many populations, the possibility of A*02 allelic heterozygotes cannot be ignored, and where the situation dictates, these additional reactions must be pursued.

The component reagents of the PCR were basically as described elsewhere in this Chapter and in Chapter 2. A negative control reaction, containing distilled water instead of DNA was included to ensure the PCR mixture was not contaminated. The first round A*02 specific amplification was performed through a 30 cycle step-down PCR. 10µl of PCR product was screened on a 2% Agarose gel stained with ethidium bromide to check for the A*02 specific amplification. A 1:100 dilution of the A*02 specific PCR product was made, and 5µl of this is added to a final 50µl volume of each of the subtyping reactions. The panel of subtyping reactions was then run through a 15 cycle step-down PCR as previously described (section 4.3.3) The subtyping reactions were then visualised on a 2% agarose ethidium bromide stained gel, and the result interpreted by the presence or absence of the appropriately sized PCR products in each of the panel reactions.

The reaction panel was tested against genomic DNA derived from a panel of sequenced HLA-A*02 B cell lymphoid lines or PBMC, which covered all the known A*02 sequences at that time (Bodmer, J.G. et al 1994; Krausa, et al 1995). The initial first round reaction (AL#37/AL#AW) was shown to specifically amplify all the known A*02 alleles. Using the A*02 specific PCR product as template DNA for the nested PCR, the subtyping reaction panel worked as expected, allowing identification of all the alleles tested. Figure 4.9 shows the patterns of reactions obtained from A*02 subtyping the first 14 A*02 alleles from the panel of control cells. Any sample typing as A*0201/A*0209 was further tested with an additional SSP combination (Table 4.11), defining the A*0209 specific polymorphism in exon 4.

Tube No.	Coding SSP	Non-coding SSP	Specificity	size
A2 FIRST ROUND				
A2	AL#37	AL#AW	A*0201-*0214	813
A2 NESTED SHORT SUBTYPING PANEL				
a	AL#22	AL#Q	A*0201, *0204, *0207, *0209, *0211	715
b	AL#22	AL#AK	A*0202	579
c	AL#22	AL#AE	A*0203	694
d	AL#22	AL#AF	A*0204	540
e	AL#3	AL#AK	A*0202, *0205, A*0214	409
f	AL#27	AL#AK	A*0205, *0208, A*0214	579
g	AL#22	AL#BJ	A*0207	549
h	AL#26	AL#AK	A*0208	408
i	AL#27	AL#Q	A*0206, *0210, A*0214	715
j	AL#27	AL#N	A*0210	546
k	AL#55	AL#Q	A*0211	522
l	AL#22	AL#R	A*0212 A*0213	705
m	AL#22	AL#BL	A*0203, A*0213	695
Distinguishing A*0209 from A*0201 (30 Cycle from genomic)				
209	AL#29	AL#AL	A*0209	907

Table 4.11. Initial A*02 subtyping panel in which an first round A*02 specific reaction is subsequently typed by an additional panel of nested reactions (a-m). A*0209 is differentiated from A*0201 by a separate reaction (209) from genomic DNA.

Additional SSP combinations were added to the subtyping panel for the determination of new A*02 alleles (A*0215N-A*0217) (data not shown). The effects of the polymorphism present within newly identified A*02 alleles on the specificity of the subtyping panel were evaluated, and additional SSP combinations included to ensure allelic A*02 definition. Specific SSP combinations for A*0216 and A*0217 were added to the nested subtyping panel. As with A*0209, a separate amplification was required for A*0215N, which identified the specific polymorphism in exon 4, differentiating it from

A*0207. An updated panel of A*02 subtyping reactions is included in Table 4.13 (section 4.5 later in this chapter).

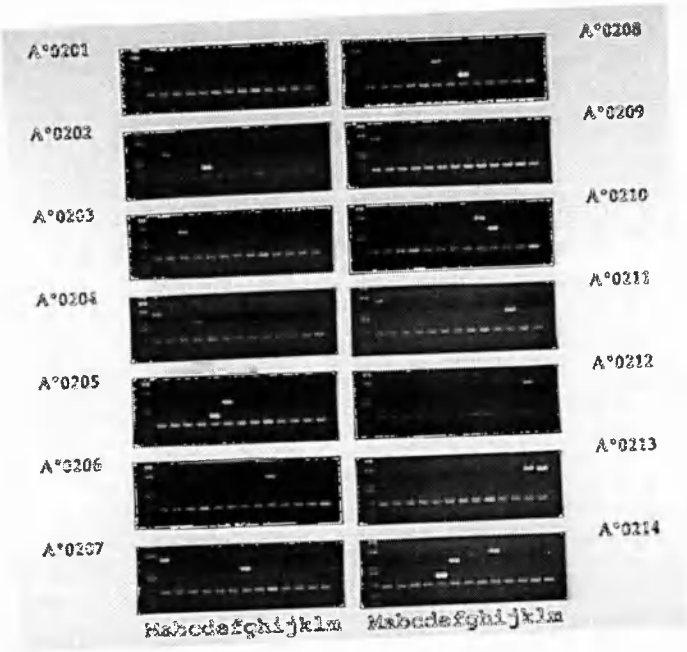


Figure 4.9. Examples of HLA-A*02 subtyping 14 sequenced DNA's using a panel of SSP combinations (a-m) as listed in Table 1. The 330bp product of the internal control primers correspond to the lower band present in each reaction lane (a-m). M= size marker. Cell lines used; JY=A*0201, M7=A*0202, DK1=A*0203, RML=A*0204, WT49=A*0205, CLA=A*0206, KNE=A*0207, KLO=A*0208, OZB=A*0209, XLI.ND=A*0210, KIME.ND=A*0211, KRC033=A*0212, SLUGEO=A*0213, 1S=A*0214. A*0201 and A*0209 were indistinguishable with this panel of nested PCR reactions but were discriminated by an additional reaction given in Table 4.11.

Reaction	Coding SSP	Non-coding SSP	Specificity	size
a	AL#22	AL#BG	A*0201/2/3/7/9/11/12/13 not A*0204	540
b	AL#22	AL#BF	A*0201/2/3/4/9/11/12/13 not A*0207	549
c	AL#14	AL#H	A*02 not A*0211	437
d	AL#27	AL#BF	A*0205/6/8/14 not A*0210	549
e	AL#22	AL#AO	A*0201/2/4/7/9/11/12 not A*0213/03	705
f	AL#27	AL#U	A*0205/A*0208 not A*0214	716
g	AL#7	AL#Q	A*0201/4/6/7/9/10/11 not A*0214	614

Table 4.12. Further reactions designed to identify additional A*02 alleles in situations that cannot exclude an A*02 heterozygote combination. (A*0201 in the presence of A*0204[a], A*0207[b] and A*0211[c], A*0206 in the presence of A*0210[d], A*0212 in the presence of A*0213[e], A*0205 in the presence of A*0214[f], and A*0206 in the presence of A*0214[g]). These reactions are necessary when A*0204, A*0207, A*0210, A*0213 and A*0214 have been identified, and no other A locus alleles have been detected. The reactions are performed nested from the first round A2 reaction. The above reaction 'f' can replace reaction 'f' in the subtyping panel (Table 4.11).

The nature of polymorphism amongst the HLA-A*02 variants occurs mainly as a result of different combinations of sequence motifs shared with other A*02 and Class I alleles at hypervariable regions within the gene. Only a few A*02 alleles contain unique sequence motifs. The lack of allele specific motifs makes the use of a nested PCR, in which the first round reaction excludes all non A*02 Class I alleles, a powerful approach for differentiating between this group of highly similar alleles. The different possible combinations of A*02 polymorphic motifs generates a potentially large number of HLA-A*02 subtypes in excess of alleles presently identified. Not all of these potential combinations of motifs will generate new alleles which offer a selective advantage to the host and result in selection of the variant within the gene pool of a given population. There remains a possibility, however, that new variants of A*02 will be found through the discrimination offered by DNA based methods of tissue typing. This is demonstrated in the characterisation of A*0214 and A*0216 discussed in Chapter 6.

The ability to identify the allelic variants of HLA-A*02 allows the potential for determination of the functional relevance of polymorphic differences, in terms of peptide binding and T cell recognition. Reports have already been made identifying different A*02 variants eliciting different profiles of T cell response (Biddison, et al 1980; Latron, et al 1991; Rotzschke, et al 1992; Utz, et al 1992; Tussey, et al 1994). It is therefore expedient to type at the highest possible level of discrimination, which in the first instance requires allelic typing to identify which sequences are immunologically relevant. Previous reports have described limited subtyping of HLA-A2 variants based on biochemical (Guttridge, et al 1992) or molecular biological approaches (Fernandez-Viña, et al 1992; Allen, et al 1994; Tiercy, et al 1994). This section represents a definitive report of the discrimination of all known HLA-A*02 alleles at the time of writing, and all possible heterozygous combinations of these alleles. Many reports have been made with regard to HLA-A2 restricted CTL responses, without identification of the A*02 allelic variant. Even in a small group of caucasians, biochemical definition of HLA-A2 may not reflect the level of discrimination of HLA-A*02 variants recognised by T lymphocytes (Gotch, et al 1985).

As discussed in Chapter 5, the diversity of A*02 alleles varies greatly in different population groups. With HLA-A*02 present at high frequencies in most populations, the method described offers the capacity to identify A*02 at an allelic level, providing cellular immunological studies with greater meaning and definition by appropriate matching of T cell and antigen presenting cell. This is particularly important when performing CTL assays in matched rather than autologous target cell lines. The capacity to identify HLA class I genes at an allelic level elucidates further population differences in HLA structure and diversity, and may increase disease associations with particular subtypes, so extending the understanding of why such differences exist.

4.5 A comprehensive PCR-SSP typing system for identification of HLA-A locus alleles.

This section concludes Chapter 4 and provides a summary of SSP combinations used either in one-step or nested PCR to achieve allelic definition at the HLA-A locus. The majority of reactions required have been described previously in this and the preceding chapters. Where appropriate, new SSP combinations acknowledge recently identified alleles, and the specificity of existing SSP combinations are modified accordingly. As many allele-specific polymorphisms have potential functional significance in terms of cellular responses (Gotch, et al 1985; van der Poel, et al 1986; Tussey, et al 1994), it becomes important to type at a level of resolution sufficient to identify these differences. The ability of PCR-SSP protocols to define HLA-A locus specificities to an allelic level holds precedent for allelic typing at other HLA loci.

The reaction combinations described in Table 4.13, are the result of development of the method through the course of this thesis. The SSP combinations listed provide the means for high resolution typing in the context of the currently identified HLA-A locus alleles (Bodmer, et al 1995). Where possible, the reactions have been screened against DNA obtained from cell lines sequenced for the relevant allele. A list of control cell line DNA tested is given in chapter 2, Table 2.1. If sequenced cell line DNA was not available, then DNA obtained from serologically well characterised cell

lines was used. In some instances, control DNA has not been available from certain recently identified HLA-A locus alleles, namely for A*0215N, A*2404, A*2405, A*2406, A*2602, A*2603, A*2604, A*1102, A*3302. The reactions identifying these alleles have therefore not been positively tested, however, they should provide the appropriate specific reaction based on the published gene sequence. Certain allele specificities are determined through reaction pattern. This has been kept to a minimum by using allele specific reactions where possible, and by the inclusion of additional reactions to ensure that certain alleles will not be hidden in particular heterozygous combinations. However, the reaction panel in Table 4.13 would not determine A*3303 in the presence of A*3301, A*1101 in the presence of A*1102, A*0204 in the presence of A*0217 and A*0207 in the presence of A*0215N.

The SSP sequences are provided in Table 2.2 and 2.3. However, to accommodate new sequence data and higher resolution typing, several of the SSP combinations listed in Table 4.13 have been altered from those previously described in this chapter and chapter 3. All SSP's have been designed to have annealing temperatures in the range of 56°C to 60°C, as a means of ensuring that all the SSP combinations work under the same conditions. These SSP's can be modified to work under conditions described by others (Olerup, et al 1993; Bunce, et al 1995a). This is can be achieved by extending the SSP sequence at the 5' end by the appropriate number of nucleotides based on the allele sequence, such that the annealing temperature of the SSP is increased to 58-62°C.

Many of the reactions can be performed in one step from genomic DNA, as shown in the subtyping of A*10 variants in figure 4.10. Typing is also possible through at least two sequential rounds of nested amplification without false positive amplification. For example, in typing A*3401 and A*3402 in sequential rounds of PCR, genomic DNA was specifically amplified with the general A locus SSP combination (CLI5'/AL#AW). This PCR product was diluted and used for a nested low resolution typing. The A10 positive PCR product (AL#38/AL#AE) from the low resolution typing was then diluted and used in a further round of nested A10 subtyping reactions.

Amplification was either directly from genomic DNA and through 30 step-down PCR cycles, or nested from a diluted first round PCR product, through 15 step-down cycles. These parameters may require alteration for use in different PCR machines. In general terms, raising or reducing the annealing temperature increases or decreases the stringency of the PCR respectively.

Amplification in 0.5ml tubes was carried out in a Hybaid Thermal Cycler, but recently PCR has been performed in the PTC-200™ thermal engine (MJ Research Inc.) which accommodates a 96 or 192 well Thermowell thin walled plate (Costar). The reaction volume used in the 96 well plate was 25µl and a 13µl reaction volume was used in the 192 plate. The use of plates increases the throughput of samples, especially when used in combination with a multitrack pipette and dedicated gel apparatus (Flowgen). As with tubes, plates containing all the reaction components except target DNA, can be stored frozen at -20°C, and thawed prior to use. Typings have been successfully performed using plates which have been stored frozen in excess of 6 months.

The use of nested PCR has proved advantageous in four main applications. Firstly, a group specific first round amplification can be used to exclude the presence of unwanted, potentially cross reactive specificities from the DNA for the nested PCR, so increasing the specificity of the second round PCR to allow for allele specific HLA gene detection. Secondly, where DNA is limiting, the use of a first round flanking PCR effectively amplifies and creates more template DNA. In this way, a full HLA-A locus typing can be performed on DNA samples as small as 100ng. In both these applications of nested PCR, first round PCR product was diluted (normally 1:100), and 1µl added per 10µl of reaction volume in the nested second round PCR. Thirdly, with the correct stringency, the first round amplification produces a highly pure template DNA, contributing to a cleaner amplification in the nested PCR. This may be advantageous for DNA samples from which one-step PCR gives relatively poor yields of specific PCR product. Finally, the use of nested PCR enables a higher resolution of typing, enabling allelic definition. Details of which subtyping reactions can be amplified from which specific first round flanking reaction are given in Table 4.13. It is important to note that a balance needs to be struck between dilution of the first round product and the number of cycles in the nested PCR amplification. Insufficient dilution

of product or too many cycles in the PCR may generate false positive amplification.

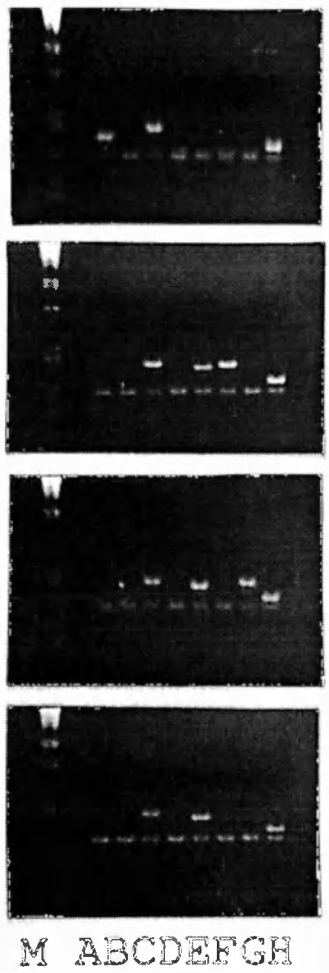


Figure 4.10. An example of one step HLA-A10 subtyping from genomic DNA, using a panel of reactions drawn from Table 4.13;

Lane A = A*25
Lane B = A*2601/2/4, A*4301
Lane C = A10 not A*4301
Lane D = A*4301
Lane E = A*34, A*66
Lane F = A*3401
Lane G = A*3402
Lane H = A10 not A*6602.

Hence, A*2501 is defined by postives with reactions A,C,H, A*3401 is defined by positives with reactions C,E,F,H, A*3402 is defined by positives with C,E,G,H and A*6601 is positive with C,E,H. M is the size marker and the lower band present in each reaction is the internal control.

One significant advantage of the PCR-SSP approach to HLA typing is that it is relatively simple to change the level of resolution by selection of SSP combinations which specifically amplify broad or narrow allele groups, making the system readily adaptable to give levels of resolution of HLA specificity to suit the application. The approach generally followed for HLA typing in studies performed in this thesis, has been to perform an initial low resolution typing (Chapter 3.2). Following this low resolution typing, further subtyping reactions were pursued as dictated by the initial result. Subtyping can be carried out nested from either first round broad specific

1st round PCR	Template		Reaction Specificity	coding SSP	non-coding SSP	band size (base pairs)
1	G	•	HLA-A	CLI5'	AL#AW	795
	G,1		A*01/A*36	AL#16	AL#AT	526
	G,1	•	A*01	AL#16	AL#Z	574
	G,1		A*0101	AL#22	AL#X	737
	G,1		A*0102	AL#12	AL#AT	671
	G,1	•	A*3601	AL#16	AL#V	563
2	G	•	A*02	AL#37	AL#AW	813
	G,1,2		A*02	AL#13	AL#H	487
	2		A*0201, A*0204, A*0207, A*0209, A*0211, **A*0215N** A*0216 A*0217	AL#22	AL#Q	715
	2		A*0202	AL#22	AL#AK	579
	2		A*0203	AL#22	AL#AE	694
	2		A*0204, A*0217	AL#22	AL#AF	540
	2		A*0202, A*0205, A*0214	AL#3	AL#AK	409
	2		A*0205, A*0208	AL#27	AL#U	716
	2		A*0205, A*0208, A*0214	AL#27	AL#AK	579
	2		A*0207 **A*0215N**	AL#22	AL#BJ	549
	2		A*0208	AL#26	AL#AK	408
	2		A*0206, *0210, A*0214	AL#27	AL#Q	715
	2		A*0210	AL#27	AL#N	546
	2		A*0211	AL#55	AL#Q	522
	2		A*0212 A*0213	AL#22	AL#R	705
	2		A*0203, A*0213	AL#22	AL#BL	695
	2		A*0216	AL#13	AL#AD	595
	2		A*0217	AL#22	AL#BQ	533
	2		A*0201/2/3/7/9/11/12/ 13/16 not A*0204/17	AL#22	AL#BG	540
	2		A*0201/2/3/4/9/11/12/ 13/16 not A*0207/**15N**/17	AL#22	AL#BF	549
	2		A*02 not A*0211	AL#14	AL#H	437
	2		A*0205/6/8/14 not A*0210	AL#27	AL#BF	549
	2		A*0201/2/4/7/9/11/12/ **15N**/16/17 not A*0213/03	AL#22	AL#AO	705
	2		A*0201/4/6/7/9/10/11/ **15N**/16/17 not A*0214	AL#7	AL#Q	614
	2		A*02 not A*0216	AL#13	AL#V	595
	G		A*0209	AL#29	AL#AL	907
	G		**A*0215N**	AL#29	AL#BS	971
3	G,1	•	A*03	AL#7	AL#D	626
	3		A*0301	AL#6	AL#AP	532
	3		A*0302	AL#6	AL#AQ	532
	G,1		A9	AL#8	AL#H	470
	G,1	•	A*2301	AL#8	AL#Q	555
	G,1	•	A*24	AL#8	AL#R	555
	G,1		A*2301/A*2402/**4/5/6**	AL#9	AL#Z	581
	G,1		A*2403	AL#8	AL#W	589
	G,1		**A*2404**	AL#2	AL#R	516
	G,1		A*2301 **A*2405**	AL#8	AL#P	521
	G,1		**A*2406**	AL#8	AL#U	558

1st round PCR	Template		Reaction Specificity	coding SSP	non-coding SSP	band size (base pairs)
4	G,1	•	A10	AL#38	AL#AE	649
	G,1,4	•	A*2501	AL#11	AL#C	398
	G,1,4	•	A*2601/2/4 A*4301	AL#34	AL#C	400
	G,1,4		**A*2602**	AL#34	AL#BN	395
	G,1,4		**A*2603**	AL#14	AL#C	407
	G,1		**A*2604**	AL#34	AL#BO	535
	G,1,4	•	A10 not A*4301	AL#4	AL#C	438
	G,1,4	•	A*4301	AL#17	AL#C	440
	G,1,4	•	A*34/A*66	AL#6	AL#C	417
	4		A*3401	AL#18	AL#C	432
	4		A*3402	AL#19	AL#AB	440
	4		A*3402	AL#4	AL#AM	426
	G,1,4		A10 not A*6602	AL#21	AL#C	362
	G,1		A*6602	AL#6	AL#AD	553
	G,1	•	A*1101, A*1102	AL#6	AL#I	518
	G,1		**A*1102**	AL#23	AL#I	675
	G,1	•	A*11, A*6601	AL#6	AL#AC	553
	G,1	•	A28	AL#6	AL#H	445
	G,1	•	A*6801 A*3402	AL#4	AL#AM	426
	G,1	•	A*6802	AL#25	AL#H	623
	G,1	•	A*6901	AL#6	AL#Y	405
5	G,1	•	A19[not A*30]	AL#38	AL#F	580
6	G,1	•	A*29	AL#35	AL#F	515
	6		A*2901	AL#17	2901	394
	6		A*2902	AL#17	AL#AS	394
	G,1	•	A*30	AL#12	AL#G	560
7	G,1	•	A*30/A*3302	AL#12	AL#W	720
	7		A*3001	AL#6	AL#K	518
	7		A*3002	AL#10	AL#T	557
	7		A*3003	AL#36	AL#T	557
	7		A*3004	AL#10	AL#U	572
	G,1,5	•	A*3101	AL#10	AL#F	481
	G,1,5	•	A*3201	AL#11	AL#F	421
	G,1,5	•	A*33	AL#4	AL#F	461
	G,1		A*3301	AL#4	AL#AG	598
	G,1,5		**A*3302**	AL#12	AL#F	598
	G,1,5	•	A*3201/A*7401	AL#24	AL#AR	494
	G,1,5	•	A19[not A32/A30]	AL#32	AL#F	414
	G,1	•	A*8001	AL#54	AL#BK	543
	G,1	•	A2/A9/A28	AL#30/31	AL#H	446
	G,1	•	A1/A36/A3/A10/A11/ A19/A80	AL#30/31	AL#L	446

Table 4.13. SSP combinations for determination of HLA-A locus alleles. Reactions listed in the '1st round PCR' column can be used as template for the appropriate subtyping reactions. The letter G in the template column denotes genomic DNA. Hence, the allele A*0101 can be amplified using SSP combination AL#22/AL#X from genomic DNA or 1st round reaction 1 product (CLI5'/AL#AW) to give a band 737 base pairs in size. A suggested HLA-A locus typing panel of reactions which can be performed directly from genomic DNA is denoted by (•). Alleles denoted in bold have not been tested.

reactions where possible (A*02 - AL#37/AL#AW, A*03 - AL#7/AL#D, A10 - AL#38/AL#AE, A19{not A30} - AL#38/AL#F), or from the A locus specific general amplification. This approach does not suit all applications, and for a more routine typing at the HLA-A locus, a panel of 31 reactions is suggested (identified in Table 4.13), which provide a level of resolution which exceeds that achievable by serology in a single round of PCR.

At the time of development, the reaction panel listed in Table 4.13 provides definition of all published A locus alleles (Bodmer, et al 1995). However, the current activity of DNA sequencing has generated a constant stream of new alleles both for class I and class II specificities. Many recently identified HLA Class I alleles appear to have arisen as the result of gene conversion events, which implies that many polymorphisms in a newly identified allele may be shared with others. This will increase the possibility that a recently identified allele will be amplified by an existing amplification, so altering the specificity of one or more reactions. The integrity of the reaction panel has therefore to be validated within the context of newly identified allelic sequences. If the specificity of an established reaction is altered by a newly identified allele, then it may be necessary to find an additional or alternative SSP combination.

The capability for such high resolution definition opens the way for an increased understanding of anthropological differences, the functional significance of allelic differences to aspects of T cell responses, their relevance to transplantation and insight into the functioning of the HLA restricted cellular immune response.

4.6 Discussion.

This chapter charts the development of PCR-SSP subtyping at the HLA-A locus from the level of resolution obtainable with serology to a level achieving allelic definition. The methods described provided the means of discrimination of all the HLA-A locus alleles known at that time. High resolution was possible through a combination of ARMS designed SSPs, stringency within the PCR, validation with analysis of control DNA, and

where appropriate, the introduction of additional specificity in the PCR through nested approaches.

The developments described in this and the previous chapter formed the basis of the A locus component of a class I PCR-SSP DNA typing kit and an A*02 subtyping kit, which was used in the recent 12th International Histocompatibility Workshop. Additionally, certain of the A locus SSP combinations have been incorporated into a class I and II PCR-SSP 'Phototyping' technique described by Bunce et al (Bunce, et al 1995b). The ability to type simultaneously for both class I and II in one procedure, with resolution beyond that determined through serology, foretell the increased use of standardised PCR-SSP and other DNA techniques in the field of histocompatibility testing.

PCR-SSP tissue typing realises specificity through the identification of two points of cross-dependent sequence polymorphisms, as determined by the specificity of each SSP in the PCR reaction. Hence, for a particular allele to be amplified, it must contain two particular polymorphic sequence motifs along the length of its sequence, for which each SSP has relevant specificity. The resulting PCR product must then be of expected length, as determined by the sites of specificity of each SSP. However, as more alleles are identified, and the complexity of the HLA system increases, it is increasingly difficult to identify allele specific reactions using one-step PCR. Nesting allows additional levels of specificity. The specificity of the first round allows the possibility of greater resolution in subsequent nested subtyping reactions. This is achieved through the specificity of the initial flanking reaction removing the presence of sequences which may interfere in determination of the allele-specific tissue-type in the subsequent subtyping reactions. Furthermore, sequential nesting can be used in determining allelic definition for certain groups of alleles. Table 4.13 outlines the possible different routes to obtaining typing at the appropriate resolution.

The ability to HLA type with increased discrimination offers the opportunity to better characterise population groups, show more precise linkage disequilibrium between HLA specificities, and better correlate HLA in disease association. The improved resolution of PCR-SSP over serology also allows improved matching in cellular assays providing more meaningful

results. Accurate determination of the tissue type, and examining the difference in function between alleles of similar sequence, may provide a better understanding of the role of polymorphism in the mechanisms of the HLA restricted immune response.

The following chapters will explore some of the above applications of this improved approach for HLA typing, with particular reference to population studies, characterisation of new alleles and studying the effects of polymorphism with regard to the functioning of the T cell mediated immune response.

Chapter 5

Analysis of HLA in populations by PCR-SSP.

5.1 Introduction.

The ability to tissue type for A locus specificities can be applied to the characterisation of HLA in different population groups. The identification of HLA polymorphism at a resolution beyond that possible through serological discrimination can increase the definition with respect to the structure of HLA gene frequencies between populations, and strengthen observations of HLA haplotype associations or linkage disequilibrium.

The highly polymorphic nature of the HLA system has provided a genetic basis for comparing different population groups. Through past workshops, based mainly on serological data (Imanishi, et al 1992a), the prevalence of HLA specificities was shown to vary considerably between different population groups. Hence HLA specificities common in one population may be rare or absent in another. The comparison of the HLA composition in different population groups may provide an understanding of how HLA polymorphisms may have evolved. The geographical distribution of populations in relation to HLA frequencies can also elucidate the origins and relationship between different populations. The presence or abundance of certain HLA specificities within a population may be in consequence to the local pathogenic pressures.

It is known through recent sequencing as reported in nomenclature reports (Bodmer, J.G. et al 1995), that a level of polymorphism exists beyond that detectable through serological characterisation. Although serology has identified clear differences between population groups, the prospect of examining differences between populations based on allelic HLA definition may emphasize the singularity of an ethnic group in comparison to the generality of serological observation. Accurate HLA characterisation of a population is also an important prerequisite for CTL based vaccine trials.

The characterisation of population groups in terms of HLA differences based on allelic differences may elucidate differences not identified previously. In this chapter, HLA-A specificities in populations will be characterised to determine which allelic subtypes are prevalent, and if this information strengthens any linkage disequilibrium. On a practical note, these exercises advocate the use of DNA methodology over serology, since exotic populations can be collected distant to the site of typing without the requirement for separating lymphocytes and keeping them viable for testing.

The studies included within this chapter involve the characterisation of HLA-A*30 subtypes in the Sardinian population, comparing the prevalence of A*02 subtypes in a number of different populations, and the use of PCR-SSP 'phototyping' (Bunce, et al 1995b) with additional allelic subtyping for characterising a Ugandan population.

5.2. Characterising the allelic variants of HLA-A*30 in a Sardinian population.

5.2.1 Introduction.

The origins of the Sardinian population, are far from clear, but the first human settlement is estimated in the Paleolithic period, at around 15,000-10,000 BC. A succession of settlements ensued during the following ages, including those from South West Europe, Asia Minor, the Aegean and Balkan regions reflecting the trade links taking place in the Mediteranean. The inclusion of several ethnic groups, such as Corsican, Balearic, Iberian, Central European, Libyan, and Estruscan has also been made into the

Sardinian population. The impact of this on the native population is not readily apparent. Also, a major civilisation was present between 2000BC - 300BC, prior to a succession of invasions, starting with the Phoenicians at around 500BC (Contu, et al 1992).

HLA characterisation therefore invited an interesting perspective upon the anthropology of the Sardininian population. This was approached in an extensive study of some 551 families, generating interesting data on the HLA structure of the population (Contu, et al 1992). The HLA analysis was performed by serological typing. As described in chapter 4, PCR-SSP offered a more discriminatory method of HLA typing, revealing polymorphism not determined through serological testing. To demonstrate this, individuals serologically identified as HLA-A30 were subtyped for this allele by nested PCR-SSP.

This initial HLA characterisation (Contu, et al 1992) observed the presence of certain haplotypes, characteristic of the Sardinian population and also showed a high incidence of HLA-A30 (23.02% frequency), mainly as part of an A30 Cw5 B18 DR3 haplotype. HLA-A*30 is one of the specificities included in the HLA-A19 serologically cross reactive group. But at a nucleotide level, HLA-A*30 seems distinct from the A19 group, and probably belongs within an A1/A3/A11 allele group (Kato, et al 1989). Prior to characterising this population, three allelic variants had been identified in the contemporary nomenclature report (Bodmer, J.G. et al 1994), with a fourth identified as a consequence of this study. These variants can be differentiated by one-dimensional isoelectric focusing (IEF) (Yang, 1989a; Choo, et al 1993), but are difficult to distinguish by serology. Hence, this study describes the use of PCR-SSP in application to A*30 subtyping. The increased level of typing definition which can be achieved using molecular techniques contributes additional information to the understanding of the HLA composition in population studies.

PCR-SSP approaches for identification of A*30 subtypes have been described in Chapter 4 (sections 4.2.5, 4.4.3 and 4.3). In this study, a nested approach was pursued as outlined in Chapter 4.4.3. The use of a nested typing offers a highly discriminatory approach to defining alleles of very similar sequence, even when they are present in heterozygous combination.

As described below, this approach has been used to subtype DNA from a panel of unrelated Sardinian individuals who had been typed by serology as HLA-A30.

5.2.2 Materials and methods.

Thirty-five serologically characterised HLA-A30 DNA samples were taken from unrelated members of the Sardinian Population study panel (Contu, et al 1992) and HLA haplotype information was available on twenty-nine of these individuals. The Sardinian samples were all at 100ng DNA per µl TE buffer. DNA samples from well characterised B cell lines, including LBF, RSH (A*3001), EJ32B, CR-B (A*3002), and JS (A*3003), were used as controls, and were at varying concentrations diluted in dH_2O. All DNA's were used at 500ng/50µl for the first round PCR reaction. The PCR-SSP nested subtyping was generally performed as previously described in Chapter 4.4.3

5.2.3 Results.

Thirty-five HLA-A*30 samples of which twenty-nine had haplotype information were subtyped to identify alleles A*3001, A*3002, and a new A*30 variant, A*3004. A*3003 was not found in the sample population. In chapter 4, Figure 4.7 shows typing results representing HLA-A*30 alleles, A*3001, A*3002, and A*3003 amplified by nested ARMS/PCR following first round amplification with primers AL#12/AL#W. A*3004 only amplified in this nested subtyping PCR panel with SSP pair AL#10/AL#O (not shown) (Table 4.10), distinguishing it from the other known A*30 alleles. This observation can be explained by the results obtained in the series of PCR gene mapping reactions discussed in chapter 6.

The haplotype results in terms of HLA-B locus, based on gene segregation in family studies, are given in Table 5.1. Additionally, six HLA-A*30 samples in which HLA gene segregation studies had not been performed, were also subtyped (not included in Table 5.1). Two of these samples typed as A*3001, and four typed as A*3002. All four A*3002 samples from this group contained HLA-B18 in their B locus assignment. Counting HLA-A*30 homozygotes as two alleles, the sample group contained thirty-three A*3002

alleles, four A*3001 alleles and one A*3004. In all, twenty-six out of thirty-three A*3002 alleles were associated with B18.

HAPLOTYPE			NUMBER
A*3002	- B18		22
	- B50		2
	- B41		1
	- B8		1
	- B49		1
	- B27		1
	- B35		1
		total	29
A*3001	- B50		1
	- B22		1
A*3004	- B49		1
		total	3

Table 5.1. HLA-A*30 haplotypes in a Sardinian population. In 32 HLA-A*30 haplotypes, 2 subtyped as A*3001, 1 subtyped as A*3004 and 29 subtyped as A*3002. 22 of the 29 A*3002 types were associated with B18.

5.2.4 Discussion.

This study has demonstrated that the use of nested PCR-SSP can be applied to defining the allelic variants of HLA-A*30 in a Sardinian population. The high frequency of HLA-A*30 in Sardinia, particularly in linkage disequilibrium with B*18, provided an interesting group in which to define the A*30 allelic subtypes. The data shows that of thirty-eight Sardinian HLA-A*30 alleles typed, thirty-three were A*3002, four were A*3001 and one a new variant, A*3004 (chapter 6.3). No A*3003 alleles were identified. Further to this, linkage disequilibrium was demonstrated between A*3002 and B18 in Sardinia, (twenty-two out of twenty-nine A*3002 haplotypes tested in this study). The A*3002, B18, Cw5, DR3 represents the most common haplotype in the Sardinian population. Linkage disequilibrium between HLA-A30 and B18 has been reported in other population groups (Contu, et al 1992), namely Basques, Spanish, Catalan and Berber populations, but no A*30 subtyping data exists for these population groups. These results do confirm previous

IEF data which suggests linkage disequilibrium between A30.1 (A*3002) and B18 (Yang, 1989a). The prevalence of A*3002 in Sardinia, contrasts with frequencies observed in other population groups. Data from another study in three African Black population groups found that in twenty-nine HLA-A*30 alleles typed, twelve were A*3001, sixteen were A*3002 and one allele typed as A*3004. As discussed in chapter 6, the characterisation of A*3004 shows that it represents the IEF variant A30.2 which has been found in low frequencies in both Caucasian and Black populations (Yang, 1989a).

In conclusion, nested PCR-SSP was used to type further serologically defined HLA-A30 samples to achieve an allelic level of discrimination. Examining the allelic nature of A*30, present in the Sardinian population at high gene frequency (23%), it can be seen that 91% of the A*30's type as A*3002 of which 69% are associated with B*18. The data obtained demonstrated the potential benefits of high resolution DNA typing in elucidating the allelic HLA structure of a population and strengthening serologically perceived linkage disequilibria. The high resolution of this typing method has given further insight into the structure of HLA amongst the Sardinian population.

5.3 The allelic variation of HLA-A*02 in different populations.

5.3.1 Introduction.

HLA-A*02 is found at high frequency in the majority of populations investigated. The gene frequency of HLA-A*02 has been defined by serology in a British population at 24%, in Black populations at 18% and in Singapore Chinese at 37% (Imanishi, et al 1992a). The prevalence of A*02 in so many populations may be explained by its capability to present a peptide motif (Falk, et al 1991) using anchor residues commonly found in many protein sequences.

The first PCR-SSP HLA-A*02 analysis of population groups was in the context of the 14 alleles then known. The number of published A*02 alleles has increased to seventeen at the time of writing (Bodmer, J.G. et al 1995). Identification of many of the A*02 variants was not supported by frequency data concerning their prevalence within the population they were identified

nor any other population group. In this study, the nested subtyping approach described in chapter 4.4.4 and 4.5 provided the means for identifying the A*02 allelic variants in a number of different populations. This involved the performance of an initial A*02 specific amplification which was used as template for a panel of PCR reactions using priming sites nested to the initial PCR.

A nested subtyping reaction panel has been used to identify the relative frequencies of A*02 alleles in a number of different ethnic groups, namely Caucasoid, African black and a Singapore Chinese population. Each of the populations showed a distinct distribution of A*02 alleles amongst the HLA-A*02 individuals investigated.

Discrimination between HLA-A*02 allelic variants allows differentiation between potentially functionally different HLA Class I molecules. This has implications for organ transplantation, in studies of T lymphocyte responses in HLA-A*02 individuals, particularly in non-Caucasoid populations, and in population genetics and disease association studies.

5.3.2 Material and Methods.

Genomic DNA samples were obtained from a number of different population groups. The HLA-A*02 populations studied included 63 DNA's from random cadaver donors from Oxford England, 55 random individuals from the Chinese Singapore population and 40 random, and 19 selected Sardinian individuals. A number of different African populations were also investigated for A*02 heterogeneity. These included 45 random individuals from The Gambia, 73 individuals from Nairobi, and 27 from Uganda. Additionally, 7 individuals from west Central Africa and 10 individuals from elsewhere in Kenya were also A*02 subtyped.

The studies were performed at different times, and the A*02 specificity in the subtyping panel reflects the known A*02 variants at the time of investigation. The reaction panels used for typing all except the Ugandan population was as described in Table 4.11 (Chapter 4). The Ugandan population was typed with a reaction panel specific for the 17 A*02 allelic variants as described in Table 4.13 (Chapter 4).

The use of the subtyping reaction panel as described in Table 4.11 may mean that examples of A*0216 may have been missed in populations typed with this reaction panel. Additionally, A*0215N may have been missed in samples typing as A*0207. A*0217 may have been missed in samples typing as A*0204, but this allele was not identified in any of the studies.

The Singapore Chinese population was also typed for HLA-B*46, to investigate if linkage disequilibrium existed with a particular A*02 allelic variant. HLA-B*4601 was identified using primer combination BL#46 (5' GAGACACAGAAGTACAAGCG 3' specific for B*4601 and all HLA-C alleles) and BL#D1 ((Sadler, et al 1994); specific for a subset of HLA-B alleles including B*4601).

5.3.3 Results

Oxford Caucasoid population.

63 random DNA samples were obtained from serologically identified HLA-A2 donors in Oxford. 13 individuals were identified as being HLA-A2 homozygous by serology and this was confirmed by a generic ARMS-PCR low resolution HLA-A locus typing (chapter 3, Table 3.5). The samples were A*02 subtyped by nested ARMS-PCR as described above and the results are shown in Table 5.2. This population predominantly contained A*0201 (96.1%), although one allele was identified as A*0205 (1.3%), one as A*0211 (1.3%) and one as A*0213 (1.3%). The A*0205 sample and one A*0211 sample were found in heterozygous combination with A*0201 and this result was validated by the presence of only HLA-A*02 in the generic low resolution HLA-A locus typing.

Singapore Chinese population.

55 Singapore Chinese DNA samples from individuals who had been identified HLA-A2 by serology, were A*02 subtyped. The sample population contained six HLA-A*02 homozygotes and five HLA-A*02 heterozygotes, which again were verified by an ARMS-PCR low resolution typing (chapter 3, Table 3.5). The profile of A*02 allelic frequency amongst this population

	Oxford Caucasian n=76	Sardinian n=44	Singapore Chinese n=66	Gambian n=47	Nairobi n=88	Uganda n=30
A*0201	96.1	59.1	22.7	57.5	56.8	63.3
A*0202	-	-	-	31.9	28.4	23.3
A*0203	-	-	22.7	-	-	-
A*0204	-	-	-	-	-	-
A*0205	1.3	40.9	-	10.6	9.1	6.7
A*0206	-	-	7.6	-	-	-
A*0207	-	-	45.5	-	-	-
A*0208	-	-	-	-	-	-
A*0209	-	-	-	-	-	-
A*0210	-	-	1.5	-	-	-
A*0211	1.3	-	-	-	-	-
A*0212	-	-	-	-	-	-
A*0213	1.3	-	-	-	-	-
A*0214	-	-	-	-	5.7	6.7
A*02 gf*	24.1	28.7	37.0	14.1	-	15.6

Table 5.2 A comparison of the proportion of the allelic variants of HLA-A*02 in five population groups.
* A*02 gene frequencies (gf%) as reported for a British, Singapore Chinese population (Imanishi, et al 1992a), and Sardinian population (Contu, et al 1992). The Ugandan population gf is taken from section 5.4. The A*02 gf of the Nairobi population has not been reported.

varied considerably from the other populations studied (Table 5.2). As a percentage of the A*02 alleles tested, A*0207 was the most frequent (45.0%), with A*0203 (23.0%), A*0201 (23.0%) and A*0206 (7.5%) also significantly represented. A single example of A*0210 (1.5%) was found in this population. Any sample which typed as A*0207 was tested for the presence of A*0201 using SSP combination AL#22/AL#BF (A*02 but not A*0207). Two A*0201/A*0207 heterozygous samples were also found. The study group also contained one A*0201/A*0210 heterozygous sample, one A*0201/A*0203 heterozygous sample, and one A*0203/A*0206 heterozygous sample.

A significant association has been described between HLA-A*02 and -B*46 in Singapore Chinese, with a haplotype frequency of 11.4% and a linkage disequilibrium value of 5.8 (Imanishi, et al 1992a). It was of interest to determine whether this linkage disequilibrium occurred with any particular A*02 allele found in the Singapore Chinese population. Of the 66 A*02 DNA alleles identified, 19 were B*46 positive when typed by ARMS PCR-SSP. Of these 19 B*46 alleles, 17 typed as A*0207 (one B*46 sample was an A*0201/A*0207 heterozygous sample). Using these data in conjunction with previously published frequency figures for HLA-A*02 in Singapore Chinese (Imanishi, et al 1992a), it would be expected that A*0207 would be found at a gene frequency of 16.7% in the Singapore Chinese population. With B*46 at a gene frequency of 15.1%, then a A*0207/B*46 haplotype would be expected at a frequency of 2.5%, if no linkage disequilibrium existed. The observation in this study was a frequency of 9.6%, suggesting a linkage disequilibrium value of 7.1 (observed - expected). This suggests that in the Singapore Chinese population, A*0207 is the main A*02 allele in linkage disequilibrium with B*46. The frequency data generated for the population in these studies are limited by the number of samples investigated and need confirmation through larger investigations.

Sardinian population.

As with the study of A*30 allelic variants (section 5.2), DNA was obtained from a well characterised Sardinian population (Contu, et al 1992). The DNA cohort, consisted of 40 random individuals, serologically defined as A*02. This groups contained 4 A*02 homozygotes individuals. 19 selected A*02

subjects were also investigated, and this group contained 5 individuals who were A*02 homozygous. The 19 selected individuals contained an A*02 B17 haplotype, commonly found in the Sardinian population. The results obtained with the randomly selected A*02 individuals, indicated the presence of two A*02 subtypes in the Sardinian population, namely A*0201 and A*0205 (Table 5.3). Of the 44 random haplotypes tested, 26 (59.1%) typed as A*0201 and 18 (40.9%) typed as A*0205. As shown in Table 5.2, this is in marked contrast to the Oxford Caucasoid population, in which A*0201 represented 96.1% of the A*02 alleles and A*0205 was rare (1.3% of A*02's tested).

HLA-A*02 individuals			Selected A*02-B17 haplotypes		
Alleles	N°	Frequency	Haplotype	N°	Frequency
A*0201	26	59.1%	A*0205-B58	20	83.3%
			A*0201-B58	3	12.5%
A*0205	18	40.9%	A*0201-B57	1	4.1%

Table 5.3. Results obtained in subtyping 44 A*02 alleles from 40 individuals and 24 selected A2-B17 haplotypes from the Sardinian population.

Because the Sardinian population was well characterised through family studies, in which the haplotype could be determined, the associations between the A*02 allele identified and the B locus allele was known and is shown in Table 5.4. The most frequent association with A*0201 at the B locus was HLA-B*18, observed in 7 of the 26 A*0201 haplotypes identified. The association at the B locus with the other A*02 allele, A*0205, was much more marked. Of the 18 alleles identified as A*0205, 13 were in linkage disequilibrium with B17. This observation is confirmed in examining the 24 selected A*02-B17 haplotypes shown in Table 5.3. Subtyping of these selected samples shows that 20 of the 24 A*02 alleles type as A*0205. All of the 20 A*0205 were associated with B*58, whilst 4 A*02's subtyped as A*0201, 3 being associated with B*58 and one associated with B*57.

If the data from the 40 random individuals is combined with the selected set, out of 40 A*02-B17 haplotypes tested, 33 (82.5%) were A*0205 and only 7 (17.5%) typed as A*0201. This shows that there is strong linkage disequilibrium between A*0205 and B17. It has previously been observed

that A2 is associated with B58 in an extended haplotype also containing Cw7, DRB1*1601, DQA1*0102, DQB1*0502 present at a frequency of 5.99% in the Sardinian population. It seems from the observations made in this study, that A*0205 is a constituent component of this haplotype.

HLA-A*0201-B haplotypes			HLA-A*0205-B haplotypes		
Haplotype	N°	Frequency	Haplotype	N°	Frequency
A*0201-B18	7	15.9%	A*0205-B17	13	29%
A*0201-B35	4	9%	A*0205-B21	2	4%
A*0201-B51	4	9%	A*0205-B35	1	2%
A*0201-B17	1	2%	A*0205-B13	1	2%
A*0201-B21	1	2%	A*0205-B39	1	2%
A*0201-B14	1	2%			
A*0201-B52	1	2%			
A*0201-B37	1	2%			
A*0201-B45	1	2%			
A*0201-B44	1	2%			
A*0201-Bbl	2	4%			

Table 5.4 . Frequency of A*02-B haplotypes obtained from 40 random A*02 Sardinian individuals.

*A*02 subtyping in African populations.*

Having analysed HLA-A*02 in both oriental and Caucasoid populations, the next major ethnic group that was examined was from Africa. Three populations were subtyped for HLA-A*02. The populations in this study were from The Gambia, Uganda and Nairobi in Kenya. The Gambia is geographically distinct from the other two populations being located in West Africa as opposed to the Uganda and Kenya, which are in East Africa.

The individuals typed were taken from cohorts being used in disease studies. The Gambian DNA was drawn from individuals taking part in either malaria (Allsopp, et al 1992; Aidoo, et al 1995), or HIV studies (Ariyoshi, et al 1996). The Nairobi DNA was taken from a study of sexually transmitted diseases in commercial sex workers (Plourde, et al 1992). The

DNA from Uganda was taken from a random HIV cohort used to characterise the HLA structure of the population. This information from the Ugandan study will be used in the design and implementation of a future cellular based vaccine trial. The typing of this cohort is discussed more extensively in section 5.4 of this chapter.

The results as indicated in Table 5.2 show a considerable similarity in the three African populations typed. All three have A*0201 as the prevalent A*02 allele (56.8-63.3%), with A*0202 also present in significant numbers (23.3-31.9%). A*0205 is again found in all three populations, but represents only a smaller proportion of A*02 population (6.7-10.6%). From the observation made in these three African populations, it seems that A*0214 is found in low numbers in East Africa (5.7-6.7%) but is absent in the Gambian population. It would be interesting to investigate the extent of A*0214 distribution, and confirm its absence in West Africa. This observation has limitations in the number of individuals typed. Both A*0202 and A*0214 seem to be restricted to the African population groups.

In addition 7 A*02 samples were typed from a West Central African population and 11 A*02 samples were typed from a study on a series of unrelated adults living in the north-western region of Kenya, in Saradidi, a town on the shores of lake Victoria. The West Central African population yielded five A*0201, and one of A*0202 and A*0205. The Saradidi population produced nine A*0201 and one A*0202 individuals, and one new A*02 variant which was sequenced and named A*0214. The characterisation of this new variant is described in the following chapter (chapter 6.3.2).

5.3.4 Discussion.

Section 5.3 describes the definition of genetic polymorphism within HLA-A*02, and the relative frequencies of individual A*02 alleles in a number of different ethnic populations. Significant differences in the range and relative frequencies of individual A*02 alleles were observed both between and within the Caucasoid, Oriental and African populations. In the African population groups and more so in the Singapore Chinese population, the diversity of A*02 alleles is more pronounced as compared to the Caucasoid group studied. A*02 heterogeneity is also observed between more closely

related populations. This is exemplified by the absence of A*0214 in the Gambian population as compared to the Ugandan and Kenyan populations and more markedly, the high frequency of A*0205 in the Sardinian population in contrast to the Oxford Caucasoid population. With HLA-A*02 present at high frequencies in most populations, application of nested PCR-SSP subtyping offers the capacity to identify A*02 at an allelic level, providing cellular immunological studies with greater meaning and definition by appropriate matching of T cell and antigen presenting cell. This is particularly important when performing CTL assays in matched rather than autologous target cell lines.

The polymorphisms described encode genetic substitutions in amino acid residues in and around the peptide binding groove of the HLA-A2 molecule. The selection and maintenance in different populations of polymorphic variants may reflect mutations which confer a selective advantage on the individual in terms of protective immunity against potentially lethal pathogens. Such a mechanism has been proposed to account for the high incidence of HLA-B53 in African populations, as this specificity has been associated with protection from severe malaria (Hill, et al 1991b) and a cytotoxic T cell response to the malaria parasite, restricted by HLA-B53 (Hill, et al 1992).

Several studies have confirmed that polymorphism between A*02 alleles is relevant for peptide presentation. Recently, the peptide motif has been determined for A*0207 (Sudo, T. et al., 1995) and the presence of an Asp P3 anchor residue makes it distinct from the other A*02 peptide motifs identified (Rotzschke, et al 1992; Barouch, et al 1995; Sudo, et al 1995). This may hold relevance for understanding the prevalence of A*0207 over A*0201 in the Singapore Chinese, which may have arisen through pathogenic pressure (Lawlor, et al 1990), such that A*0207 individuals may have obtained a selective advantage within the Singapore Chinese population. Hence the heterogeneity of HLA-A*02 variants across different ethnic groups may be a consequence of the variation of pathogenic pressures acting upon those populations.

The polymorphism present within different A*02 variants have implications for matching donor and recipient for transplantation. As

different population groups contain different A*02 variants, which elicit different immune responses, then it would be prudent to know to which ethnic group recipient and donor belong. As can be seen in the data generated in this study, trying to match A*02 between a Caucasoid and a Singapore Chinese individual cannot rely on serological or biochemical definition. Preliminary evidence from allelic typing in unrelated bone marrow transplant screening show a significant incidence of mismatch for HLA-A*02 alleles between potential donors and recipients which have been typed by serology and/or one dimensional isoelectric focusing (Anholts et al, personal communication). For this reason, HLA population studies which can identify allelic frequencies, may prove essential for recognising which variants to match for in a given ethnic group. Also, linkage disequilibrium, as described between B*46 and A*0207 in the Singapore Chinese A*02 samples and A*0205 and B*58 in the Sardinian population, becomes much more apparent with a high resolution typing method, and may provide useful information through haplotype association in situations where high resolution typing is not available. With PCR-SSP offering allelic typing, further cases of linkage disequilibrium may be found, which remain obscured or diluted by the relatively poor level of resolution offered by serology.

This study shows a marked variation in the relative frequencies of A*02 alleles in different populations. It also reveals a difference in A*02 allelic diversity within each population. This diversity may reflect the pathogenic pressures each population has been subjected to. As the capability to determine the allelic composition is applied to an increasing number of population groups, so a better understanding of the structure and diversity of HLA polymorphism will follow. This may determine more succinct characterisation of ethnicity which in turn may better chart the evolution of HLA polymorphism through the study of different populations. The ability to define allelic variants could strengthen disease associations with particular subtypes and extend our understanding of why such differences exist.

5.4 HLA Class I & II polymorphism detected by PCR-SSP in a semi-urban HIV positive Ugandan population.

5.4.1 Introduction

The study of HLA polymorphism with regard to different population groups has revealed widely disparate results in terms of the HLA specificities found and their relative frequencies (Imanishi, et al 1992a). Until recently, determination of tissue types in population studies has mainly relied on serology for Class I (Ferrone, et al 1973) and a number of DNA based approaches for Class II (Erlich, et al 1991a; Vaughan, 1991; Olerup, et al 1992). In characterisation of certain populations, particular HLA specificities prove difficult to identify by serology, due to a lack of relevant specific antisera or misleading crossreactivity. Alternative DNA based methods now exist, which provide an accurate and easily interpreted approach to HLA Class I and Class II typing (Bunce, et al 1995b).

The purpose of this study was to identify the HLA Class I and Class II polymorphism in a typical group of HIV infected Ugandans. These individuals will be used to identify the characteristic HLA restricted cytotoxic T lymphocytes (CTL) directed against HIV viral epitopes present in the population. Similar individuals will be participating in a forthcoming HIV vaccine trial, based upon the initial CTL work. Reports have suggested that some individuals have developed partial or complete protection against the HIV virus and this may in part be afforded to HIV specific T cells (Rowland-Jones, et al 1993; Gallimore, et al 1995; Rowland-Jones, et al 1995). Such reports provide the rationale for the development of a possible prophylactic vaccine which could stimulate cellular immune responses and in particular virus specific CTL, thus protecting the individual against HIV infection (McMichael, et al 1994).

CTL exhibit specificity through HLA Class I restriction and the sequence and orientation of foreign peptide bound by the HLA molecule (McMichael, et al 1988; Madden, et al 1993; Rammensee, et al 1995). The HLA restriction of CTL clones has been demonstrated to differentiate beyond that determined by serology (Gotch, et al 1987; Tussey, et al 1994; Barouch, et al 1995). Therefore, in contemplating such a vaccine trial study, it is important to know the frequency and fine specificity of HLA alleles within a particular

population, in conjunction with the clades of virus circulating within that population. Certain CTL will exhibit cross reactivity between viral strains, however differences in viral sequence known to exist in Uganda could mean that CTL responses in Sub-Saharan Africa will be to totally different to those already identified in Caucasoid populations, even if restricted by the same HLA specificity.

In comparisons between African and Caucasoid populations, it is immediately apparent that major differences exist in terms of the gene frequencies found between both groups (Figure 5.1). Differences can be seen even between different African populations (Figure 5.2) and tribal groups, (duToit, et al 1988; duToit, et al 1990; Allsopp, et al 1992; Imanishi, et al 1992a; Halle, et al 1994; Dieye, et al 1996). For the purpose of a vaccine trial, it becomes implicit to accurately characterise the HLA polymorphism within a population. Identifying those common alleles which cover in excess of 90% of the population, provides the necessary information from which the appropriate viral epitopes can be incorporated into a vaccine inducing HLA restricted CTL.

5.4.2 Materials & Methods.

With the above considerations, 3ml of EDTA blood was collected from 90 HIV seropositive individuals attending the TASO and UVRI clinics in a relatively high prevalence area around Entebbe. Using a Puregene™ DNA Isolation kit (Gentra Systems Inc), genomic DNA was extracted from this blood using a simple salting out technique (Buffone, et al 1985; Miller, et al 1988). This yielded approximately 50-150µg DNA per 3ml whole blood. This genomic DNA was initially tested by a PCR-SSP panel of 144 'Phototype' reactions (Bunce, et al 1995b) which identifies specificities at HLA-A, -B, -C, -DRB1, DRB3, DRB4, DRB5 and DQB1 loci. Details of the Phototyping protocol is given in chapter 2 (2.1.4).

The level of resolution of the 'phototyping' panel was equivalent and in areas superior to that offered by serology, without the problems of crossreactivity, difficult heterozygous combinations or detection of HLA-Cw specificities. It also permitted the simultaneous typing for both HLA class I and class II under a standard set of conditions in a relatively short period of

time. In application to a non-Caucasoid population, this PCR-SSP allowed clear identification of certain HLA specificities to which serological reagents may be rare or ambiguous.

In view of the population group studied, additional allele specific PCR-SSP subtyping was performed for certain prevalent HLA group specificities. Hence, HLA subtyping was performed on samples identified as HLA-A*02, -A*29, -A*30. -A*68 (chapter 4, Table 4.13) and B*58 (Browning, et al 1995). The results of the A*02 subtyping are described in chapter 5.3.4 of this chapter. B*5801 and B*5802 could be discriminated in the Phototyping panel except when in the presence of B*5301. It was important to identify these subtypes correctly since they are prevalent in this population and can potentially elicit functionally distinct cytotoxic T cell responses (Barouch, et al 1995). Determining allelic subtypes involved the use of nested PCR (chapter 4.4). This required amplifying an initial flanking PCR reaction which contained the relevant sites of polymorphism necessary for discriminating the allelic group of interest. This first round product was then used as template DNA for a subsequent nested panel of PCR reactions which define the relevant subtypes (chapter 4).

5.4.3 Results.

Table 5.5 shows the gene frequencies defined by 'Phototyping' in this Ugandan population with those alleles derived by additional subtyping reactions highlighted. In the Ugandan population, of the 90 individuals, 27 typed as HLA-A*02 and following nested subtyping (chapter 4.4.4 & 4.5), it was revealed that the gene frequencies (gf) of each variant was A*0201 (9.94%), A*0202 (3.41%), A*0205 (1.11%) and A*0214 (1.11%). Gene frequencies (gf) were determined by the maximum likelihood method (Imanishi, et al 1991a). These data again shows differences in frequency profiles as compared to other population groups in previous studies (section 5.3). HLA-A*30 proved the most frequent HLA-A locus specificity detected amongst the Ugandans typed. Nested subtyping (chapter 4.4.3 & 4.5) of this group of 31 HLA-A*30 individuals, found gene frequencies of A*3002 (9.94%), A*3001 (8.13%) and A*3004 (1.66%). In subtyping the sixteen A28 and fifteen A*29 positive individuals, all were found to be A*6802 and A*2902 respectively.

HLA-A	gf(%) (n=180)
A*0201	***9.94***
A*3002	***9.94***
A*6802	***9.34***
A*2902	***8.73***
A*3001	***8.13***
A*7401	6.89
A*2301	6.89
A*01	5.71
A*03	5.71
A*3601	5.71
A*0202	***3.41***
A*3402	3.41
A*6601	2.84
A*33	2.84
A*3004	***1.66***
A*32	1.66
A*31	1.11
A*0205	***1.11***
A*0214	***1.11***
A*26	0.55
A*6602	0.55
A*80	0.55

HLA-B	gf(%) (n=180)
B*53	11.17
B*1503	8.73
B*42	8.13
B*5802	***7.48***
B*57	5.71
B*45	5.71
B*81	5.13
B*5801	***5.13***
B*07	3.98
B*18	3.41
B*51	3.41
B*49	3.41
B*1510	2.84
B*08	2.22
B*1401	2.22
B*35	2.22
B*13	1.66
B*44	1.66
B*46	1.66
B*48	1.11
B*1516/17	1.11
B*1402	1.11
B*27	1.11
B*38	0.55
B*39	0.55
B*41	0.55
B*50	0.55
B*56	0.55
B*1501/4-8	0.55
B*67	0.55
B70v	0.55
B*73	0.55
B*1509	0.55
B*1513	0.55

HLA-C	gf(%) (n=180)
Cw*04	23.97
Cw*06	15.68
Cw*17	12.42
Cw*0701	11.80
Cw*02	9.94
Cw*0302/4	6.89
Cw*16	4.55
Cw*0702	3.41
Cw*08	2.84
Cw*0704	2.84
Cw*14	2.22
Cw*15	1.11
Cw*07	1.11
Cw*03	0.55
Cw*1203	0.55

HLA-DR	gf(%) (n=178)
DRB1*11	18.57
DRB1*13	15.86
DRB1*15	12.59
DRB1*0302	12.59
DRB1*07	10.06
DRB1*01	7.59
DRB1*0301/4	7.21
DRB1*12	5.77
DRB1*08	4.03
DRB1*04	1.71
DRB1*09	1.71
DRB1*10	0.55
DRB1*16	0.55

DRB3	72.15
DRB4	14.62
DRB5	13.22

HLA-DQ	gf(%) (n=176)
DQB1*06	25.37
DQB1*05	21.64
DQB1*02	19.50
DQB1*0301/4	16.04
DQB1*04	9.56
DQB1*0302	0.55

Table 5.5. HLA Class I and Class II gene frequencies (%) obtained in typing a cohort of 90 HIV positive Ugandan individuals. Gene frequencies (gf) were determined by the maximum likelihood method [Imanishi 1992]. Typings were only possible in 79 and 78 individuals for HLA-DRB1 and -DQB1 respectively. Typings were performed by PCR-SSP using a 144 reaction phototyping panel [Bunce et al 1995]. Results underlined denote those alleles determined or confirmed by additional PCR-SSP subtyping reactions.

At the HLA-B locus, certain key prevalent were again identified specificities in this population, namely B*58 (12.61%), B*5301 (11.17%) and B*1503 (8.73%). Subtyping was performed on the B*58 samples (Browning, et al 1995) to reveal that B*5802 gene frequency (7.48%) was slightly more frequent as compared to B*5801 (5.13%). The advantage of a PCR-SSP approach was underlined at the C locus, in which typing was clear and unambiguous, with serological C locus blanks (Cw*12-Cw*17) easily determined. The main HLA-C specificities identified were HLA-Cw*04 (23.97%), -Cw*06 (15.68%), -Cw*17 (12.42%) and -Cw*0701 (11.80%). The prevalent HLA Class II gene frequencies for this Ugandan population were DRB1*11 (18.57%), DRB1*13 (15.86%), DRB1*15 (12.59%) and DRB1*0302 (12.59%). At HLA-DQ, we found DQB1*06 (25.37%), DQB1*05 (21.64%), DQB1*02 (19.50%) and DQB1*0301/4 (16.04%).

5.4.4 Discussion

A comparison of broad specificities at HLA-A and -B between the Ugandan population characterised by PCR-SSP and several other serologically typed population groups is depicted in Figures 5.1 and 5.2. In figure 5.1, the Ugandan population is compared to a British and a Singapore Chinese population (Imanishi, 1992). It is clear that the prevalence and gene frequency profile of HLA specificities varies greatly between the three population groups. At HLA-A (Figure 5.1a), HLA-A*30 is the most prevalent specificity in the Ugandan population, but relatively infrequent in the British and Singapore Chinese population groups. HLA-A*02 is present at high frequencies in all three populations, however previous studies have shown that the allelic composition varies in each of the population groups (chapter 5.3). The British population is characterised by the prevalence of HLA-A*01 and -A*03, whilst the Singapore Chinese contain HLA-A*11 and -A*24 at high gene frequency.

At HLA-B (Figure 5.1b), the gene frequency profiles in the three populations were clearly different. The Ugandan population contained a high proportion of HLA-B70 (B*1503, B*1510), -B*58 and -B*5301. This is in contrast to the British population, in which HLA-B*08, -B*44 and -B*35 were the predominant specificities. The Singapore Chinese population is again different containing a high number of HLA-B*46, -B*40 and -B5 (B*51, B*52)

specificities. Many other differences exist between the three populations at HLA-A and -B, and these are represented in Figure 5.1.

Figure 5.1a

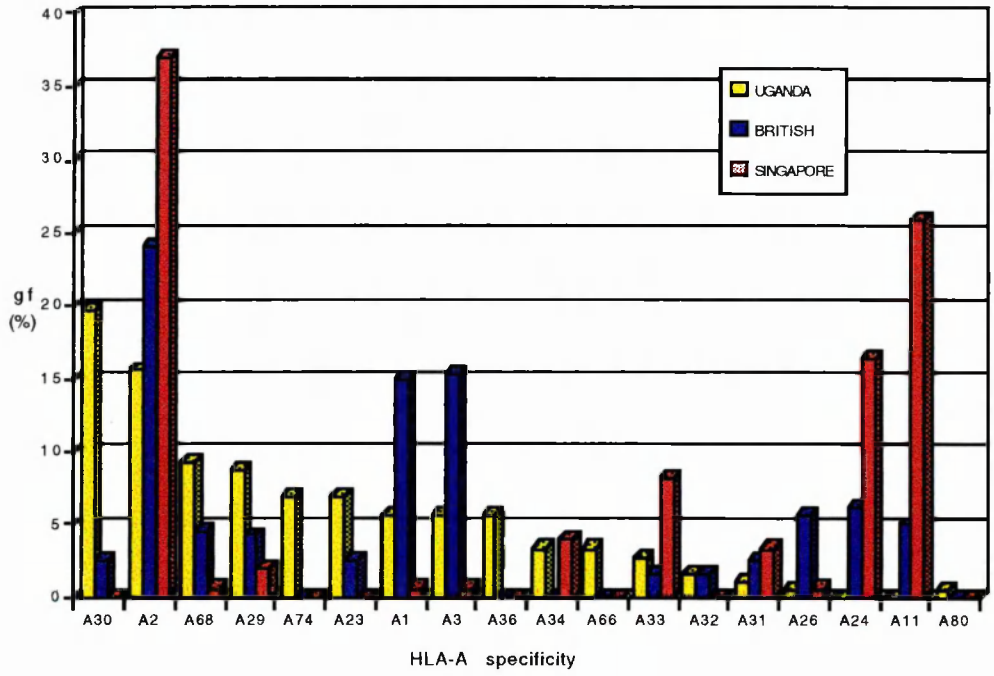


Figure 5.1b

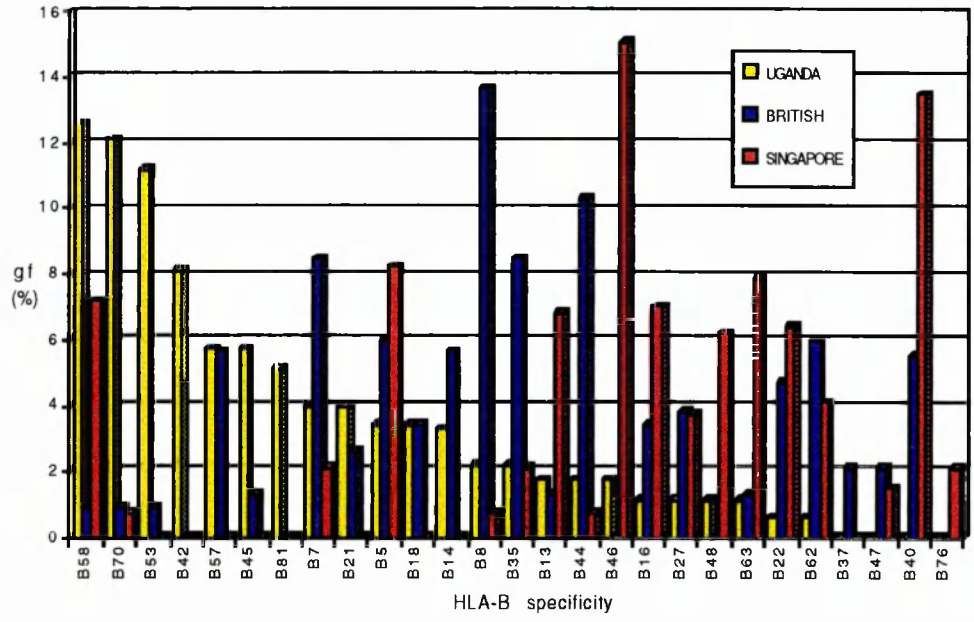


Figure 5.1. Gene frequencies observed in serologically characterised British and Singapore Chinese populations (Imanishi, 1992) in comparison with the PCR-SSP typed Ugandan study population. Frequencies were compared on the basis of serological specificities; (a) represents a comparison at HLA-A and (b) a comparison at HLA-B.

Figure 5.2a

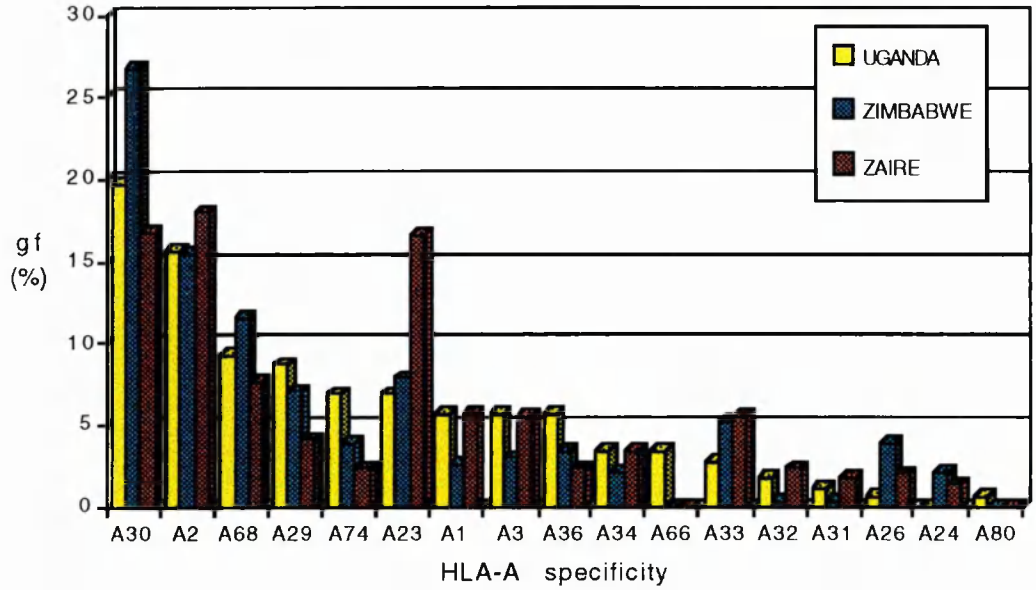


Figure 5.2b

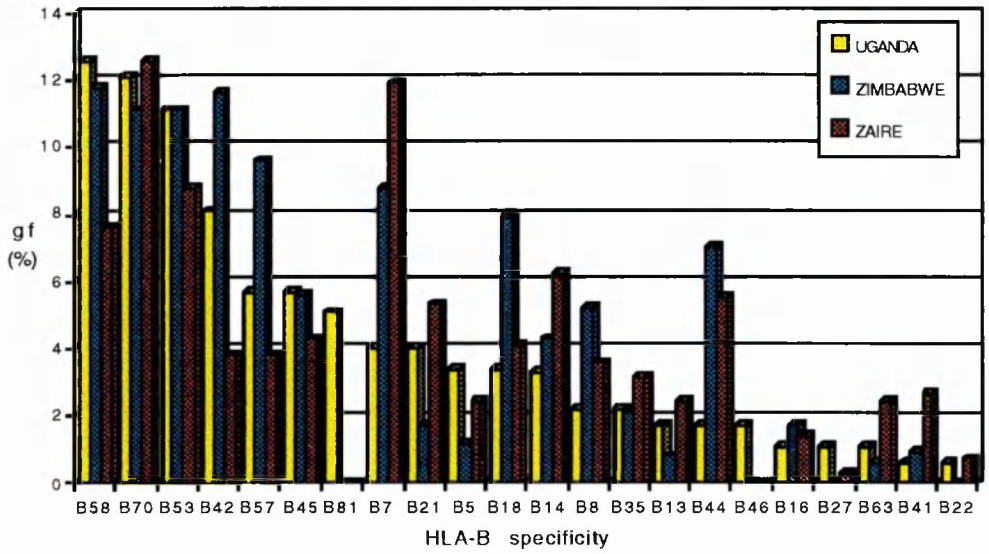


Figure 5.2. Gene frequencies observed in serologically characterised Zimbabwean and Zairean populations (Allsopp, 1992; Imanishi, 1992; Halle, 1994) in comparison to the PCR-SSP typed Ugandan study population. Frequencies were compared on the basis of serological specificities; (a) represents a comparison at HLA-A and (b) a comparison at HLA-B.

Figure 5.2 represents a comparison of the PCR-SSP typed Ugandan population in this study with serologically determined gene frequencies observed in two other African populations (Allsopp, 1992; Imanishi, 1992; Halle, 1994). Although the Ugandan population shows similarities with the two African populations as compared to the British or Singapore Chinese

populations (Figure 5.1), distinctions can still be made. Hence in both Uganda and Zimbabwe, HLA-A*30 followed by -A*02 represent the most prevalent specificities. This is reversed in the Zaire population. The prevalence of HLA-A*23 and -A*33 in both Zimbabwe and Zaire is in contrast to the lower frequency observed in the Ugandan population. At HLA-B, similarities were seen between the three populations, through the prevalence of B70 (B*1503, B*1510). However differences were also observed, as illustrated by the high gene frequency of HLA-B*58 and B*5301 in the Ugandan and Zimbabwean populations as compared to the Zairean population. It was also noted that HLA-B*57 and -B*18 were more prevalent in Zimbabwe than in the other two populations, whilst HLA-B*07 and -B*44 were found at a higher gene frequency in Zimbabwe and Zaire as compared to Uganda. Certain recently identified specificities (HLA-A*8001, HLA-B*8101) were not tested for in population groups used for comparison. It is probable that re-screening would identify HLA-A*8001 and B*8101 in some of the population groups considered. It must also be acknowledged that this population may not prove wholly typical of a general Ugandan population since it has been selected from HIV positive individuals attending two clinics.

The results generated through this PCR-SSP approach, represent a more comprehensive description of HLA in an African population than previously described. The observations in general are in concordance with those reported in a general Black population (Imanishi, et al 1992a). It should be noted that substantive differences exist between this population and Caucasoid populations in which immunological responses to HIV have been investigated. Thus in the Caucasoid population >90% of the population may be represented as having A*0101, A*0201, A*03, B*07, B*08 or B*44, whilst a different set of specificities represents >90% of the Ugandan population studied (A*0201, A*3001, A*3002, B*5301, B*5802 or B*1503). This emphasises that the main HLA restrictive elements for HIV specific CTL are in the most part likely to be different for each population. It was also noted that the frequency of B*35 which has been associated with bad prognosis in HIV infection (Itescu, et al 1995), was much lower in the Ugandan population (2.2%) as compared to West African populations (Gambia 16.1%, Senegal 13.1%) (Allsopp, et al 1992; Imanishi, et al 1992a; Dieye, et al 1996).

With the key HLA specificities identified in this study, putative CTL epitopes can be ascertained from the predominant HIV circulating gag, pol and nef proteins. The preferred peptide binding motifs of certain HLA specificities (e.g. A*3001 and A*3002) had not been demonstrated at the time of this population study. Because of their prevalence in this and other populations, the peptide motifs of the allelic variants of A*30 was determined as described in chapter 7. Identifying peptide epitopes for all the major HLA specificities in the Ugandan population, will allow strategies to be designed for the pursuit of possible HIV vaccines.

It therefore seems probable that a prophylactic vaccine, designed to stimulate potentially protective cellular immune responses, specifically in an African population will contain a totally different profile of viral peptide epitopes as compared to a vaccine specifically designed for an at risk Caucasoid population. This does not necessarily imply that a vaccine will have to be 'tailor made' for an Ugandan population, but rather a vaccine designed for global use should take into account the very different genetic background of different populations, in addition to the heterogeneity of the circulating HIV viruses.

This study therefore describes the use of PCR-SSP to HLA type a cohort of Ugandan HIV positive individuals. The level of resolution possible by this method addresses the discrimination possible by the T cell in cellular immune responses. This accurate determination of HLA types forms an important prerequisite to potential approaches in vaccine trials based on the HLA restricted immune response.

5.5 Discussion on application of PCR-SSP to population studies.

This chapter describes studies in which the PCR-SSP approach to HLA typing has been applied to analysis of HLA-A allelic variants in a number of different population groups. Characterisation of HLA has been an important aspect of anthropological studies by virtue of the polymorphic nature of the HLA system. Population analysis has also provided a certain amount of insight as to the evolution of HLA polymorphism through geographical distribution.

The 11th International Histocompatibility Workshop included an anthropological component. This component analysed the relationship between different population groups on the basis of HLA polymorphism (Imanishi, et al 1992a). The relationship was explored through the generation of dendrograms; phylogenetic trees which show an estimation of the genetic distance between different populations. Analysis revealed and confirmed the clustering of different populations into distinct ethnic groupings. However, this representation did have flaws as described in the characterisation of two Amerindian groups as being distantly related, when a good deal of other evidence shows the opposite. Further analysis was required to make a more valid judgement as to the relationship between different populations.

The analysis of genetic distance in the above workshop studies was based upon serological definition for class I specificities and some DNA based typing at class II. By including different subtypes found in different populations in one serologically defined basket, the similarities or differences between populations could be lost. This was demonstrated by the identification of B*2703 in the Gambian population (Hill et al. 1991a). This allele is identified as the most common subtype in the Gambian population and absent in most non-black population groups. By not defining such differences, serology fails to provide a precise representation of a population that takes into account key allelic HLA markers. This is further revealed by the failure of serology in identifying HLA-C specificities, for which the application of DNA approaches has provided a highly definitive method (Bunce, et al 1994a; Bunce, et al 1994b).

The application of DNA based typing technology has improved the resolution of tissue-typing and so provided a more detailed account of populations in terms of their HLA composition. Because viable cells were not a requirement for DNA based methods of typing, more populations became accessible to HLA characterisation. This and the general activity in DNA sequencing has identified an ever increasing number of HLA alleles which in turn provide more definition of ethnic groups. As more populations are studied with the tools of DNA methodology, so it seems likely that new HLA alleles will be discovered.

This has certainly been the case in the studies described in this chapter, with A*3004 being identified in the Sardinian population and A*0214 being identified in the Kenyan population. Both of these alleles were discovered through anomalies observed in PCR-SSP typing, where the sequence polymorphism they contained produced novel reaction patterns. Their characterisation will be discussed in the following chapter. The identification of A*0214 provided a marker for the East African populations, even though it was observed at relatively low frequency, since it appears absent in the West African Gambian population. It would also be interesting to examine a wider number of populations to see the true extent of the distribution of A*0214 and the other A*02 alleles.

The characterisation of A*02 in different populations, has provided a good example as to the importance of subtyping in uncovering significant differences between populations. Different alleles are present at different frequencies in different populations. Also certain alleles can be definitive of a particular ethnic group, such as A*0203, A*0206, A*0207, A*0210 in the oriental populations, A*0202 and A*0214 in African populations. This information provides a better description of the population which would be missed through serology.

Allelic definition in populations also enables a better distinction of linkage disequilibrium. Hence, through these studies, it is clear that A*3002 is in linkage disequilibrium with B*18, and A*0205 is in linkage disequilibrium with B*58 in the Sardinian population, and A*0207 is in linkage disequilibrium with B*46 in the Singapore Chinese population. Linkage disequilibrium had been noted though broad serological definition, but allelic definition both identified the precise nature of this association and so strengthened the observation. Identification of linkage disequilibria, which may vary in different populations, provides another index for characterisation of a population group.

The ability of PCR-SSP for typing a remote population is exemplified in the Ugandan study. The application of the 'Phototyping' panel (Bunce, 1995b), with additional subtyping provided a well characterised population. This information is not only useful for anthropology, but also for clinical studies.

As mentioned in the Ugandan study, this information is extremely important for the designing and implementing of a cellular based vaccine trial, in which HLA restricted responses form a major consideration. In addition, disease association can be made more apparent when subtyping is performed. Reports have been made, as to the association of A*0206 with Graves' disease and A*0207 with Hashimoto's thyroiditis, in the Japanese population. An association has also been recently established between A*6802 and susceptibility to scarring trachoma in a Gambian population (Conway, et al 1996).

PCR-SSP is therefore a powerful means of characterising a population in terms of its HLA polymorphism. The application of such definitive typing will undoubtedly provide a great deal more information about populations and so increase understanding of the relationship between ethnic groups. The characterisation of populations may also clarify the nature of HLA evolution in terms of genetic drift and selective pressure. The precise characterisation of HLA polymorphism within different population groups has an important influence in the fields of transplantation and cellular base vaccine programs. These reasons validate the importance of accurate characterisation of HLA polymorphism in different populations

Chapter 6

Analysis and characterisation of new HLA alleles.

6.1 Introduction.

Through the course of this thesis, novel HLA-A specificities have been identified either through anomalies observed in PCR-SSP typing, or discrepancies between different typing approaches. The application of a new typing approach produces a novel perspective from which the HLA alleles can be defined. By altering the means of definition, there remains a possibility that this will expose new polymorphic differences between alleles previously considered to be identical. However, PCR-SSP does not define all the polymorphism within an allele, and so there remains that possibility that new alleles will not be discriminated.

The characterisation of HLA polymorphism has relied on such discriminatory observations, be it through the differing specificity of sera or T cells, IEF and RFLP analysis, or more recently through PCR based analysis of DNA. The observation and discovery of the highly polymorphic nature of HLA specificities has been possible through the use of increasingly specific reagents, or the contradictory results obtained by two or more methods of analysis. Indeed, the development and application of new technology over the years has been instrumental in the definition of the nature and level of polymorphism within the HLA system.

The characterisation of HLA polymorphism has been further enhanced

through the comparison of different population groups. The application of methods designed for analysis of HLA polymorphism in one population may produce unexpected results when applied to different population groups. This is due to the potential presence of as yet unknown HLA specificities in uncharacterised populations. As mentioned in Chapter 5, HLA polymorphism in different populations may be driven by different pressures, which may not only shape the frequency of HLA specificities, but may also determine novel specificities unique to that particular ethnic group.

The ability to analyse HLA in different populations has increased as the restrictions of serological typing have been replaced by the more convenient and practical PCR based DNA analysis. The combination of HLA typing in different population groups together with the application of definitive DNA based typing approaches has extended the number of known alleles (Domena, et al 1993; Hildebrand, et al 1994a; Hildebrand, et al 1994b; Selvakumar, et al 1995; Date, et al 1996). This has been the case in the application of methods developed in this thesis to the study of HLA-A specificities in different populations. The study of Caucasoid and African populations has yielded a number of newly identified HLA-A alleles, identified through discrepencies in the PCR-SSP typing approaches described in previous chapters. A new HLA specificity was also identified through a discrepancy between IEF data and the PCR-SSP typing result.

To define accurately the sequence polymorphism defining these new alleles, it is necessary to characterise them by DNA cloning and sequencing. As an interim measure, to determine quickly the composition of polymorphism found within an allele, panels of ARMS designed SSP can be used to define the presence or absence of specific polymorphic residues or motifs at particular sites within the allelic sequence. This development of this PCR approach of 'gene mapping' (see section 6.2), has proved useful on a number of occasions in characterising and confirming the combinations of polymorphism denoting a new allele. This approach has also been useful in providing a quick and easy method of confirming the presence of absence of specific polymorphic motifs at specific locations within the HLA sequence. It does not represent an alternative to DNA sequencing which is still necessary for characterisation of new HLA alleles.

This chapter describes the circumstances under which new alleles were identified in this project, and the approaches taken for their further characterisation. The application of DNA based methods to HLA typing has shown a larger level of polymorphism than observed through previous serological approaches. This chapter will also discuss the relevance of polymorphisms found within the newly identified alleles, not only on the interpretation of the PCR-SSP typing approach at HLA-A locus, but also on the possible effects on immunological function. In addition, this chapter will describe the use of PCR-SSP as a means of characterising HLA sequences, either to confirm the polymorphism found within a sequence, or to ascertain the polymorphic content of a new allele. Studies are also described in which the nature of serological HLA-A 'blank' specificities are investigated. The use of PCR-SSP not only contributes to the detection of new HLA specificities, but provides the means for their characterisation.

6.2 PCR 'gene-mapping.'

As an extension of the PCR-SSP method, a technique of PCR 'gene-mapping' has been developed (Krausa, et al. 1995d) for the definition of sequence polymorphism found within HLA alleles. PCR 'gene-mapping' uses the specificity of a panel of SSP's, based on the known polymorphic motifs at specific sites within the HLA class I genes, to define the sequence polymorphisms found within a particular allele. This has proved highly useful in the evaluation of sequence polymorphism in HLA alleles. This approach of using PCR-SSP to ascertain the polymorphic 'content' of HLA specificities has been beneficial in both confirmation of known alleles, as well as the characterisation of new allelic sequences. This method of sequence characterisation is made possible through the specificity of ARMS PCR (Newton, et al 1989). ARMS PCR under the correct conditions of stringency, has the ability of discriminating between sequences on the basis of a single base-pair difference. This approach provides a highly powerful analytical tool for the detection of sequence polymorphism.

The nature of polymorphism found amongst HLA alleles provides another consideration in their identification. Very few HLA sequence

polymorphisms are unique to a particular allele, rather they are shared with other specificities not only within the same locus, but on occasions with alleles of other loci. The definition of a HLA allele is therefore mainly through the unique combination of these sequence polymorphisms along the length of the gene. By having a substantive panel of SSP defining the major sequence polymorphisms found amongst HLA alleles, it is possible to identify the key polymorphisms present within a particular HLA specificity. This information can then be extrapolated to predict the DNA sequence of a newly identified allele, prior to its characterisation by formal sequencing methods (Sanger, et al 1977).

In HLA class I, the majority of sequence polymorphisms are found within exons 2 and 3 of the HLA genes, which code for the domains which form the the peptide binding groove of the molecule. This part of the molecule plays an important role in antigen presentation and interaction with the T cell. In developing the PCR-SSP for HLA typing at A locus, a panel of SSPs were assembled which defined many of the polymorphisms found within these exons. These SSP have proved useful reagents for the characterisation of new allelic sequences. The approach of PCR 'gene-mapping' for definition of polymorphism in exons 2 and 3, is to combine a fixed specific SSP in one exon with a panel of SSPs defining polymorphisms in the other exon. The polymorphism defined by the fixed SSP provides the context within which the sequence is characterised. It allows a single allele to be analysed as opposed to SSOP in which the PCR generally amplifies both alleles. The use of nested PCR can be used to increase specificity of the SSPs in the performance of PCR gene-mapping, so providing a more definitive characterisation.

6.3 New alleles identified during the work for this thesis.

6.3.1 HLA-A*3004

HLA typing for the allelic variants of HLA-A*30 in the Sardinian population (Chapter 5.2) identified an individual, who although serologically typed as A*30, A*11, failed to amplify with a putative A*30, A*3302 specific SSP combination (AL#12/AL#Q), which had already been positively tested

against the known A*3001, A*3002 and A*3003 alleles. Although this individual was shown to be negative with the AL#12/AL#Q combination, typing with the AL#12/AL#G SSP pair confirmed the presence of an A*30 gene. The discrepant results with the serological typing, and the PCR-SSP observations, indicated the presence of a new A*30 allele.

The above observations already provided some information on the nature of the polymorphic difference between this A*30 and the other sequenced A*30 alleles. The failure to amplify with SSP combination AL#12/AL#Q, but positive amplification with AL#12/AL#G, indicated this new variant contained the polymorphisms defined by SSP AL#12 and AL#G, but not that defined by SSP AL#Q. To investigate these observations further, this variant of A*30 was characterised through PCR 'gene-mapping', with particular emphasis on the region identified by SSP AL#Q.

The A*30 variant allele (A*3004) was characterised by PCR 'gene-mapping' through a nested PCR approach using a first round product obtained with SSP pair AL#12/AL#W (specific for A*30 A*3302), which gave a PCR product of 719bp. The SSP listed in Table 6.1, in their appropriate combinations, were used to identify the sequence motifs present in the known HLA-A*30 alleles and those in the variant, A*3004, in a nested PCR system using template generated by the first round reaction using primer combination AL#12/AL#W. An initial screen showed the sequence of A*3004 differed from the other A*30 alleles in a region roughly between nucleotide positions 170-200 of exon 3. Using a panel of SSPs in a nested PCR, a fine map of the sequence motifs present in A*3004 were defined, in comparison to A*3001 and A*3002 in this region. Exon 2 was characterised using a panel of seven motif specific SSPs run off an SSP (AL#G) common to all A*30's in exon 3. Exon 3 was more extensively characterised using a panel of twenty-seven motif specific SSPs run off a fixed primer common to A*3001 and A*3002 (AL#10) in exon 2.

Results of PCR 'gene-mapping' of HLA-A*3004, at the SSP sites tested, revealed that the sequence of the allele resembled A*3002 except for a region between nucleotides position 180-195 in exon 3. The mapping data are shown in Table 6.1, and summarised on a gel in figure 6.1. This figure identifies the different motifs distinguishing A*3001 and A*3002 from

A*3004, at primer sites AL#AX to AL#U. The figure also shows the difference in exon 3 between A*3001 and A*3002 at primer sites AL#K and AL#T. A more extensive results panel is given in Table 6.1. The sequence polymorphisms identified in A*3004 between nucleotide positions 180-195 are similar to those of A*0202, A*0205, and the A*68 sequences. The sequences of the A*30 allelic group and the A*0202, A*0205 A*68 sequence group are identical in the region 3' of position 195 in exon 3. This finding suggests a gene conversion event in exon 3. The proposed sequence of A*3004 in this region based on PCR 'gene-mapping' of polymorphic motifs is given in figure 6.2, and has been confirmed by DNA sequencing as part of this study.

In addition, observations obtained by PCR 'gene-mapping' (Table 6.1a), made it clear that the cell line LBF, from which the A*3001 allele was sequenced (Kato, et al 1989), has a cytosine and not a thymidine at nucleotide position 25 in exon 2, as previously reported. This has been substantiated by sequencing LBF in the 5' region of exon 2, and confirmed by others (Olerup, et al 1994). The correct polymorphism at position 25 is present in the other A*30 alleles together with the A9 group specificity, all of which share a Serine at position 9 in the alpha-1 domain of the class I molecule. PCR 'gene-mapping' therefore has a particular application in rapid confirmation of published sequences.

Table 6.1a. Exon 2 gene mapping (for table legend, see following page).

SSP	position of SSP 3' nucleotide	Sequence motif identified	A*3001	A*3002	A*3004
AL#5	25	ct**C**	+	+	+
AL#22	25	ct**T**	-	-	-
AL#12	53	**A**g**T**gg**A**	+	+	+
AL#10	169	**A**ggcc**T**	+	+	+
AL#6	208	ca**G**	+	-	-
AL#44	227	**T**g**G**	+	-	-
AL#45	227	aga	-	+	+

Table 6.1b. Exon 3 PCR gene mapping.

SSP	position of SSP 3' nucleotide	Sequence motif identified	A*3001	A*3002	A*3004
AL#AB	18	**TA**a	+	+	+
AL#Y	48	**T**gg	-	-	-
AL#AK	58	**G**cg	-	-	-
AL#O	68	t**G**a**A**	+	+	+
AL#G	71	**A**cag**C**	+	+	+
AL#C	80	**T**ta	-	-	-
AL#F	104	**T**tg	-	-	-
AL#H	109	**A**ga	-	-	-
AL#L	109	cga	+	+	+
AL#AU	145	**A**gc	-	-	-
AL#AJ	153	**C**ca	-	-	-
AL#P	158	cag	+	+	+
AL#S	158	**A**ag	-	-	-
AL#AZ	162	**A**ca	-	-	-
AL#AE	173	**A**cg	-	-	-
AL#AX	180	**A**tgt	-	-	+
AL#J	180	gtg	-	-	-
AL#AY	182	gtg	-	-	+
AL#K	182	**TG**g	+	-	-
AL#T	182	**CG**g	-	+	-
AL#AO	183	tgg	-	-	+
AL#I	183	**C**gg	-	-	-
AL#AN	183	**A**gg	-	-	-
AL#AA	183	**A**gg	-	-	-
AL#Q	194	**T**tg	+	+	-
AL#AV	194	**TG**g	-	-	+
AL#U	195	**G**ga	-	-	+
AL#W	226	gtg	+	+	+

Table 6.1. showing details of PCR gene mapping of HLA-A*30 alleles in exons 2 (Table 6.1a) and 3 (Table 6.1b). Uppercase letters denote polymorphisms from the consensus sequence as detailed in the 1995 sequence alignments (Arnett, K.L. et al., 1995). Sequence motifs as identified by the 3' residues of SSP's listed and sequences given at positions as denoted in the 1992 sequence alignments. Reaction patterns of the SSP's against A*3001, A*3002 and A*3004 are given when combined with AL#G to map exon 2 and AL#10 to map exon 3 in a nested PCR from first round product amplified with SSP combination AL#12/AL#W.

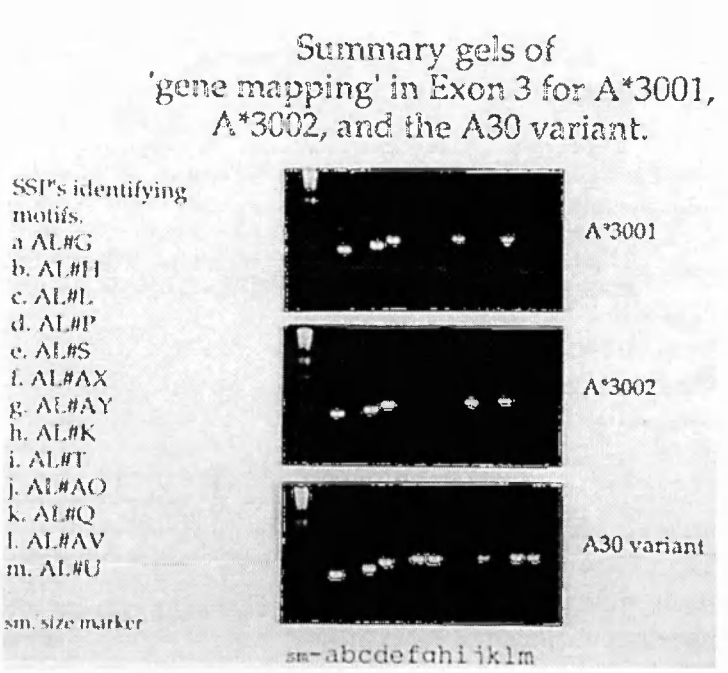


Figure 6.1. Examples of reactions for PCR gene mapping A*3001, A*3002 and the A*30 variant (A*3004) in exon 3. SSP's (a-m) defining motifs were combined with an exon 2 A*3001 A*3002 A*3004 specific SSP (AL#10) in a series of nested reactions against the AL#12/AL#W first round PCR product. A*3001 and A*3002 differ in sequence as defined by their respective reactions with SSP AL#K (h) and AL#T (i). A*3004 differs from A*3001 and A*3002 as defined with SSPs AL#AX to AL#U (f-m). A more comprehensive list of motif defining reactions is given in Table 6.1.

Sequencing A*3004 (single stranded method as described in chapter 2.2), confirmed the predictions made by 'PCR gene mapping'. In exon 2, the sequence of A*3004 is shown to be identical to alleles A*3002 and A*3003.

However, in exon 3, A*3004 is the same as HLA-A*3002 except for a region between nucleotide positions 180 and 196. At position 180, the A*3004 sequence substitutes an A for G as compared to the other A*30 alleles. At nucleotide position 182 and 183, A*3004 has bases GT as opposed to TG in A*3001 and CG in A*3002 and A*3003. At nucleotide position 196, A*3004 differs from the other A*30 alleles by having a G in substitution for a T. Details of these nucleotide differences are shown in figure 6.2. The sequence polymorphisms observed cause amino acid substitutions as compared to the other A*30 alleles. In detail, these are a substitution of histidine for arginine at amino acid position 151, valine for tryptophan (A*3001) or arginine (A*3002/3) at position 152 and tryptophan for leucine at position 156. When

these changes are viewed in terms of their location on the class I molecule (Bjorkman, et al 1987), they are all found on one of the alpha helices, pointing into the peptide binding groove. These polymorphisms have probable implications as to the peptide motifs the A*3004 molecule will bind, and hence have relevance for transplantation matching. A report of a one amino acid difference between A*3002 and A*3003 has been implicated as causing GVHD in an unrelated donor marrow transplant (Choo, et al 1993).

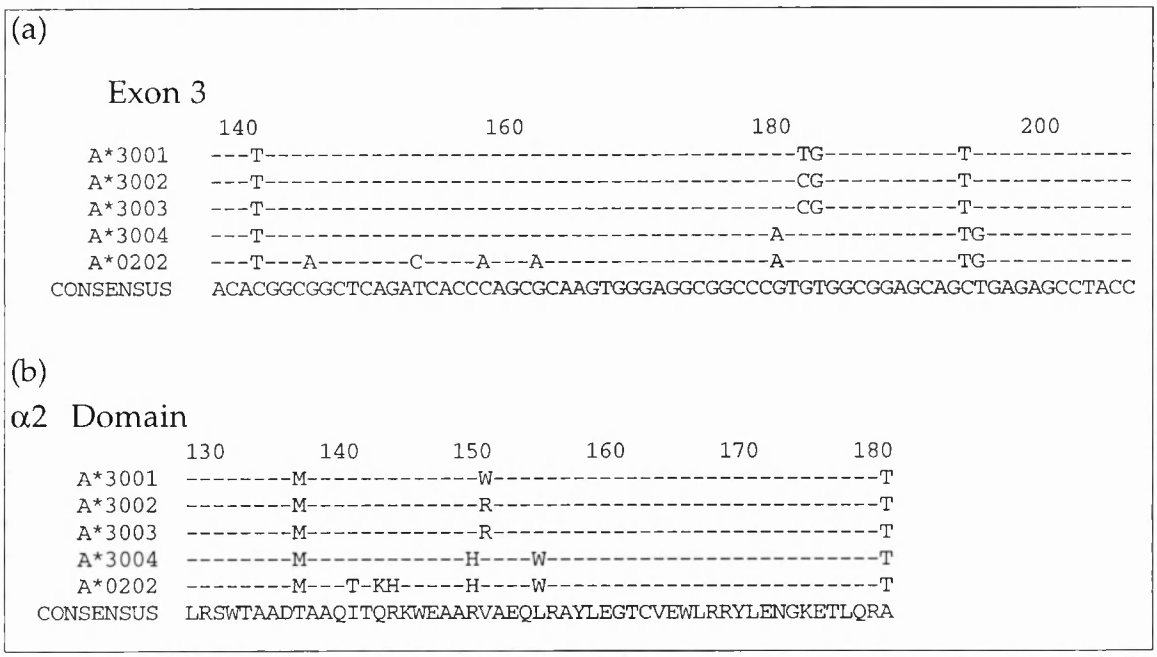


Figure 6.2. A comparison of the (a) nucleotide and (b) amino acid sequences of the variant A*30 allele (A*3004) with the other A*30 alleles, and the A*0202 sequence, showing the polymorphic motifs in the 3' region of (a) exon 3 (b) alpha-2 domain. The A*3004 polymorphisms were identified firstly by 'PCR gene mapping' and subsequently confirmed by DNA sequencing.

Further characterisation on a B cell line generated from the A*3004 individual (provided by C.Carcassi), was performed by IEF. An A*30 IEF variant had previously been described, but not sequenced (Yang, 1989a). IEF analysis was performed on A*3004 to determine whether it was the equivalent to this unsequenced IEF variant A30.2. Analysis of sequenced A*30 cell lines (LBF, A*3001; CR-B, A*3002; JS, A*3003) together with the A*30 variant (W7), were carried out as previously described (Yang, 1989b) with details given in chapter 2. Unfortunately, it was not possible to obtain

an A30.2 IEF variant control, which could provide a direct comparison with the A*3004.

Analysis of A*3004 by one dimensional IEF, as illustrated in figure 6.3, shows that the A*30 variant runs in the same position as the A30.2 IEF variant in relation to the other A*30 alleles, based on the published data of Yang et al (Yang, 1989). IEF variant A30.2 runs acidic to IEF variant A30.1 (A*3002) and basic to A30.3 (A*3001). The only amino acid differences between A*3002 and A*3004, lie in the alpha 2 domain between positions 151-156 (Figure 6.2). In this region, A*3002 has two basic residues (Arg-151, Arg-152) as compared to the aromatic and aliphatic residues (His-151, Val-152), in A*3004. This explains the relative IEF focusing of the two allele products.

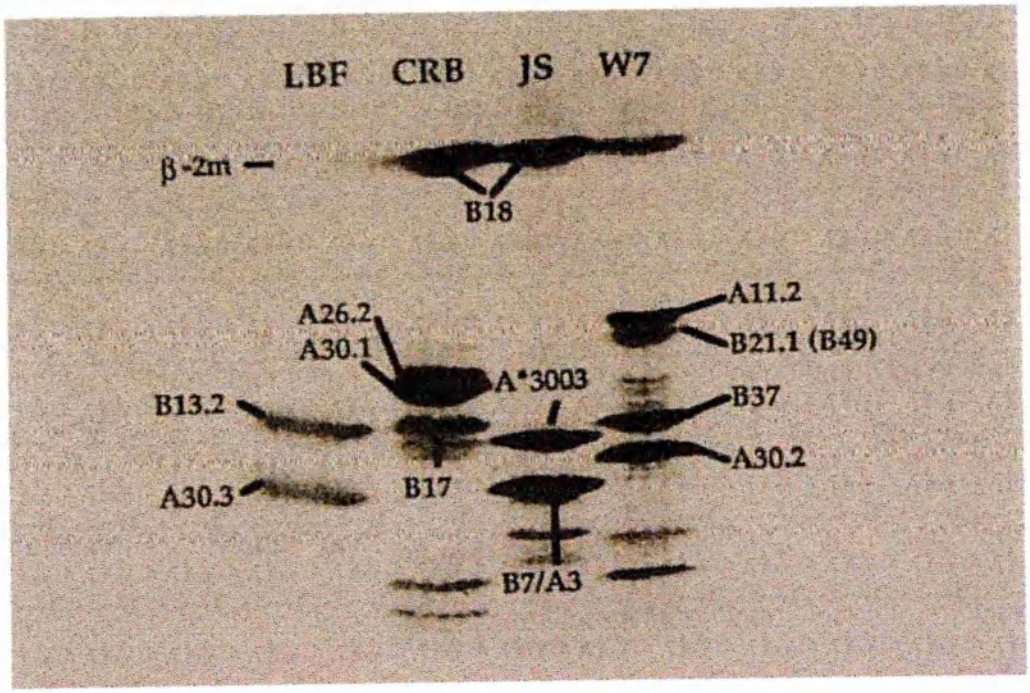


Figure 6.3. A one dimensionsal IEF comparison of four different HLA-A30 subtypes. The four cell lines LBF (A*3001), CR-B (A*3002), JS (A*3003) and W7 (A*3004) produce different A30 IEF bands. The results show LBF has an A30.3 IEF subtype and CR-B has an A30.1 IEF subtype. W7 produces an A30.2 equivalent IEF subtype. JS has an A30 IEF subtype different to A30.1, A30.2 and A30.3.

The identification of new HLA Class I alleles provides a constant reason for vigilance in terms of how new sequences may affect the specificity of existing PCR-SSP reactions. For typing A*30 alleles, we are now aware of two recently

identified alleles, A*0102 (Browning, et al 1995) and A*3302 (Kato, et al 1993), for which SSP AL#12 (previously thought to be A*30 specific) also has specificity. Of these two alleles, A*3302 can potentially be co-amplified by the first round primer combination of AL#12/AL#W. In this study and in the A*30 subtyping reactions (chapter 4.4.3), the potential presence of A*3302 does not compromise the specificity of the panel, as both HLA-A alleles were known and the presence of A*3302 excluded.

In conclusion, a presence of a novel HLA-A*30 specificity was identified through failure to amplify a serologically determined A*30 individual with SSP combinations specific for the then published A*30 alleles. The sample was typed as A*30 by PCR-SSP through the use of an alternative A*30 specific SSP combination. Additional characterisation through PCR 'gene-mapping' identified polymorphic differences to the known A*30 alleles in the latter region of exon 3. This was confirmed by conventional DNA sequencing, with the polymorphic differences identified suggestive of a probable gene conversion event by virtue of their presence in the sequences of A*0202, A*0205 A*68 alleles.

Further characterisation by IEF analysis identified the new specificity as IEF variant A30.2. This new specificity has been observed at low frequency in other population groups (chapter 5.4). The sequence of A*3004 was officially assigned as A*3004 by the WHO nomenclature committee. The nucleotide sequence is reported in the EMBL database under accession number Z34921.

6.3.2. HLA-A*0214.

As with the identification of A*3004, A*0214 was determined through an anomaly in PCR-SSP typing. A Black Kenyan individual was typed at HLA-A locus by PCR-SSP, as previously described. Genomic DNA from this individual was analysed by PCR-SSP low resolution typing, high resolution subtyping, and PCR gene mapping. Low resolution typing identified the sample as HLA-A*02, A*33. High resolution HLA-A*02 subtyping (chapter 4.4.4), suggested the individual was heterozygous for A*02, with both A*0205 and A*0206 being identified. In view of the low resolution HLA-A locus typing, this raised the possibility that the A*02 gene was a new HLA-A*02

variant which contained sequence motifs characteristic of both the A*0205 and A*0206 alleles.

PCR gene mapping of this A*02 allele (A*0214) was performed nested from the A*02 specific AL#37/AL#AW SSP combination used for nested A*02 subtyping (chapter 4.4.4). Table 6.2 lists the primers used for the identification of sequence polymorphism present in exon 2 and exon 3 of the A*0214 allele. The PCR gene mapping result (Table 6.2 & Figure 6.4) shows A*0214 to be identical to A*0205 in exon 2, but varying in exon 3 as defined by primers AL#Q, AL#U, and AL#AV, suggesting a sequence difference of one nucleotide (G→T) at position 196 in exon 3, as compared to A*0205.

exon 2	primer	primer sequence	exon 3	primer	primer sequence
First Round PCR					
	AL#37	CCTCGTCCCCAGGCTCT		AL#AW	TGGCCCCTGGTACCCGT
a	AL#22	CACTCCATGAGGTATTTCTT	a	AL#AF	ACGTCGCAGCCATACATCA
b	AL#27	CACTCCATGAGGTATTTCTA	b	AL#N	CCCCACGTCGCAGCCAA
c	AL#51	TCCATGAGGTATTTCTTCACA	c	AL#AK	ACTGGTGGTACCCGCGC
d	AL#38	AGCCCCGCTTCATCGCC	d	AL#B	CGTCGTAGGCGTACTGGT
e	AL#56	GACAGCGACGCCGCGACCCG	e	AL#AU	GTGCTTGGTGGTCTGAGCT
f	AL#40	GTCCGGAGTATTGGGACG	f	AL#AJ	CCCACTTGTGCTTGGTGG
g	AL#3	GACGGGGAGACACGGAAA	g	AL#AZ	GGCCGCCTCCCACTTGT
h	AL#26	GACGGGGAGACACGGAAT	h	AL#AE	TCCGCCTCATGGGCCGT
i	AL#6	CGGAATGTGAAGGCCCAG	i	AL#R	CTCCAGGTAGGCTCTCTG
j	AL#14	AGGCCCACTCACAGACTC	j	AL#Q	CTCCAGGTAGGCTCTCAA
k	AL#32	CGAGTGGACCTGGGGAC	k	AL#U	CTCCAGGTAGGCTCTCC
l	AL#33	GACCTGGGGACCCTGCG	l	AL#AV	CTCCAGGTAGGCTCTCCA
fixed	AL#H	CAAGAGCGCAGGTCCTCT	fixed	AL#13	TGGATAGAGCAGGAGGGT

Table 6.2. HLA specific oligonucleotide primers and sequences used to PCR 'gene-map' HLA-A*0214 in exons 2 and 3. A comparison of A*0201, A*0205 and A*0214 is shown in Figure 6.4

Confirmation of the predicted sequence (figure 6.5) was obtained by DNA DNA sequencing of exons 2-5 of the A*0214 allele as described in chapter 2.2. Single stranded DNA was prepared from single plaques of M13 phage and dideoxy sequencing (Sanger, et al 1977) was performed on both strands of three individual clones from three independent PCR amplifications.

The sequenced DNA indicated A*0214 to be identical to A*0205 in exons 2-5, except for a single base difference at position 196 in exon 3 (T instead of G), as seen in the A*0201, A*0204, A*0206-7, A*0209-11 and A*0215N-17 sequences (Arnett, 1995). In comparing A*0214 to A*0205, this causes a single amino

acid substitution of Trp_{156} (A*0205) to Leu_{156} (A*0214) in the alpha 2 domain of the molecule (Table 6.3). This change of aromatic to aliphatic side chain change at residue 156 has likely implications in the binding characteristics of the D and E pockets in the peptide binding groove (Saper, et al 1991).

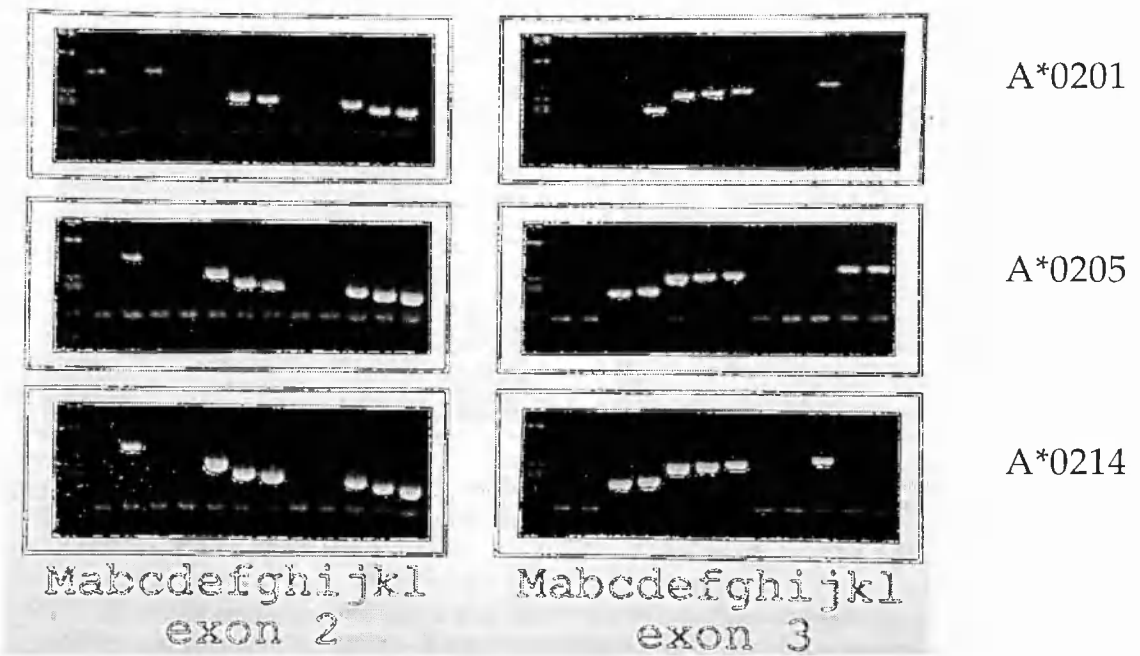


Figure 6.4. PCR gene maps of exons 2 and 3 of A*0201, A*0205, A*0214. Exon 2 gene maps were constructed using fixed primer AL#H in exon 3, and, in order, SSP AL#22, AL#27, AL#51, AL#38, AL#56, AL#40, AL#3, AL#26, AL#6, AL#14, AL#32, AL#33. Exon 3 gene maps were constructed using fixed primer AL#13 in exon 2, and, in order, AL#AF, AL#N, AL#AK, AL#B, AL#AU, AL#AJ, AL#AZ, AL#AE, AL#R, AL#Q, AL#U, AL#AV.

Exon 3

```
          160                 180                 200
A*0201    ---A-----------------A-------------T-------
A*0205    ---A-----------------A-------------TG------
A*0214    ---A-----------------A-------------T-------
CONSENSUS AGCGCAAGTGGGAGGCGGCCCGTGTGGCGGAGCAGCTGAGAGC
```

Figure 6.5. Comparison of *A*0201, A*0205* and *A*0214* nucleotide sequence, showing the only nucleotide difference in exons 2-5 between *A*0205* (G) and *A*0214* (T) at position 196 in exon 3.

A*0214 has since been identified in other Kenyan individuals and also in Uganda (chapter 5.3). However, although identified in East Africa, its

presence has so far not been noted in the West African population of Gambia, or indeed outside of the Africa. Additional A*02 subtyping of population groups may well reveal the true distribution of this A*02 variant.

Residue Position	A*0201	A*0205	A*0214
9	Phe	Tyr	Tyr
43	Gln	Arg	Arg
95	Val	Leu	Leu
156	Leu	Trp	Leu

Table 6.3. Amino acid sequence differences between A*0201, A*0205, A*0214. Note that A*0205 has four differences from A*0201 and that A*0214 has three of these four differences.

The name A*0214 has been officially assigned by the WHO Nomenclature Committee. The nucleotide sequence data reported appears in the EMBL Databases under the accession number Z30341.

6.3.3. HLA-A*0216.

The characterisation of A*0216 was possible through observations made with the B cell lymphoblastoid line TUBO (10th International Histocompatibility Workshop ID 9045; HLA-A2, A3, B51, Bw4, Cw7, DR11, DR12, DQ7, DR52, Caucasoid origin). As mentioned previously in this chapter, the discepancies observed between methods of HLA typing often highlight the presence of novel specificities. In the case of A*0216, the discrepancies presented as a difference on IEF analysis and molecular typing.

The evidence of this discrepancy exists in the initial DNA typing of TUBO by both PCR-SSOP (Fernandez-Viña, et al 1992) and nested A*02 PCR-SSP subtyping (chapter 4.4.4). Early versions of these typing approaches both typed TUBO as A*0201. However, this was at variance with the IEF analysis of the TUBO A*02. The isoelectric focal point of the TUBO A*02 was shown to be more acidic than A2.1 (A*0201), focusing at a point similar to A2.3 (A*0203, A*0209) (Guttridge, et al 1992). PCR-SSP typing of TUBO had excluded A*0203 and A*0209. The difference in the observations made in IEF

analysis and molecular typing suggested the presence of a new A*02 variant.

Verification of these observations was made in cloning and DNA sequencing (Sanger, et al 1977) of the A*02 cDNA obtained from TUBO, as described in chapter 2.2. As in the characterisation of all new HLA alleles, three clones obtained from three independent PCR amplifications were sequenced on both strands, to allow confidence in the sequence data generated. The DNA sequence data generated showed A*0216 to have a two base pair difference from A*0201, at nucleotide positions 216 and 217 in exon 3 (A*0201 AC→GA A*0216). The sequence substition caused a single amino acid change of glutamic acid instead of the threonine found at residue 163 in the alpha-2 domain of the A*0201 molecule.

The sequencing provided information for the design of an A*0216 specific PCR-SSP combination, as illustrated in figure 6.6. The observed polymorphism is not found in any other of the A*02 alleles so far described, which makes its definition by A*02 nested subtyping unequivocal, within the context of the known A*02 alleles. The polymorphism is seen in other HLA-A locus alleles, (A*6602, A*8001), in addition to a large number of B locus alleles (including B*27, B*07, B*4001-4, B*4006, B*1301-2, B*4701 & B*4801) and some C locus alleles (Cw*0201 and Cw*0202). The polymorphism being present as two adjacent base pairs, found in many other alleles, is suggestive of A*0216 arising through a gene conversion event (Parham, et al 1988). The single amino acid difference caused by this polymorphism in relation to A*0201 produces a substitution from a polar to acidic residue. This change explains previous IEF observations in that A*0216 focusing at a more acidic position in relation to A*0201.

The amino acid substitution also has implications for potential peptide binding characteristics of the A*0216 molecule. Residue position 163 of the class I molecule is known to have involvement in the binding characteristics of the 'A' pocket of the peptide binding groove (Bjorkman, et al 1987). Characterisation of this allele shows further evidence of the concentration of polymorphism around the peptide binding groove of HLA molecules, and suggests that A*0216 may be functionally distinct in terms of peptide presentation.

Further subtyping studies for A*02 has not revealed an additional example of A*0216. The name A*0216 was assigned by the WHO nomenclature committee, with the nucleotide sequence placed in the EMBL database under accession number Z46633.

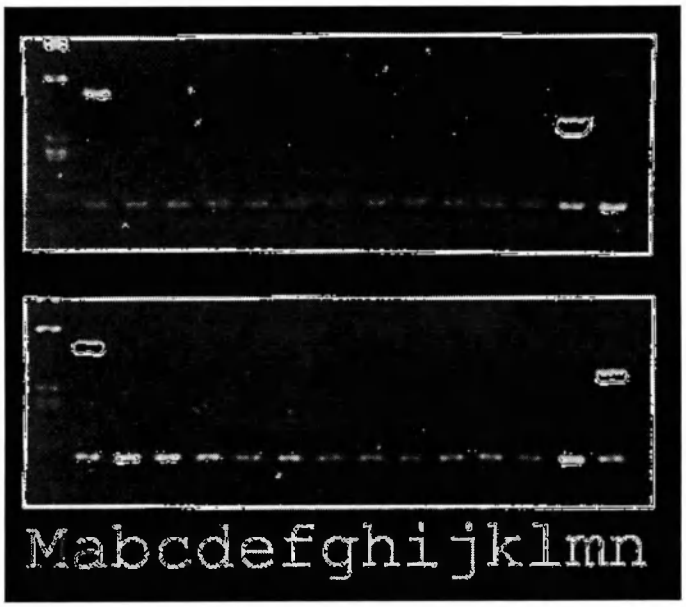


Figure 6.6 HLA-A*02 subtyping DNA from JY (A*0201) and TUBO (A*0216) by ARMS-PCR. Reaction specificity; a=A*0201/4/7/9/11, b=A*0202, c=A*0203, d=A*0204, e=A*0202/5/14, f=A*0205/8/14, g=A*0207, h=A*0208, i=A*0206/10/14, j=A*0211, k=A*0212/13, l=A*0203/13, m=A*02 not A*0216, n=A*0216. M=size marker. Lower band present in each reaction lane is the internal control. The results show the difference between A*0201 and A*0216 as determined by reactions (m) and (n) identifying the sequence polymorphism in exon 3 of the HLA-A locus gene.

6.3.4. HLA-A*2502.

HLA-A*2502 was determined through observing a novel pattern of reactivity when HLA typing a DNA sample by PCR-SSP. As discussed previously, new specificities may appear in typing, when a novel combination of polymorphisms in the allele, manifests as a new pattern of reactivity with the typing panel.

A Caucasian individual was HLA typed from genomic DNA through 'Phototyping' (Bunce, et al 1995b). The result indicated a discrepancy at the HLA-A locus, with identification of three alleles (A*6601 A*2501, A*68). The rest of the genotype comprised B*0801, B*1402, Bw6; Cw*0701, Cw*0802; DRB1*0301/4, DRB1*1303, DRB3; DQB1*02, DQB1*0301/4. Both A*6601 and A*2501 belong to a broad serological specificity, HLA-A10. It was conceivable

that the typing result represented a novel sequence recombination involving elements of A*6601 and A*2501.

To investigate further, the nature of the A*2502 sequence polymorphism was determined through PCR 'gene-mapping' as illustrated in figure 6.7 and Table 6.4. This approach found that the new allele contains a combination of polymorphisms found in both A*2501 and A*6601. The variant resembles A*2501 by containing the characteristic Bw4 motif (figure 6.8, nucleotide position 300-320), whilst it also contains a polymorphism at nucleotide position 282 found in A*6601 but not in A*2501. The combination of polymorphisms present within the sequence of A*2502 explained the initial anomalous PCR-SSP HLA typing result.

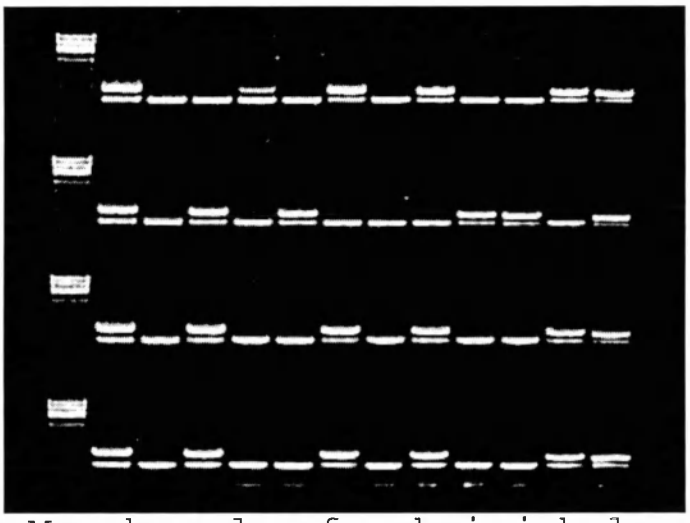


Figure 6.7 PCR gene-map of A*2502 in the 3' region of exon 2. The pattern of reactivity is compared to A*2501 (FB11) and A*6601 (TEM). Details of the gene-map panel are given in Table 6.5. Reactions were run nested from AL#38/AL#AC and constructed with a fixed primer, AL#AE, in exon 3. The lower band present in each reaction lane represents the internal control amplification.

The HLA-A*2502 gene was then cloned and sequenced, as described in chapter 2.2. Cells were then transformed and clones picked and screened by specific HLA-A PCR-SSP reactions to identify vectors containing HLA-A*2502. Four independent A*2502 clones were then sequenced on both strands. One clone was dideoxy sequenced using a standard approach (Sanger, et al 1977). Three other clones were full length sequenced on both strands using the ABI377 DNA Sequencer technology (Perkin-Elmer, USA).

Automated sequencing was performed by Angela Ivison at the Advanced Biotechnology Centre, Charing Cross & Westminster Medical School, London. DNA sequencing verified the results observed in the initial PCR-SSP tissue-typing and 'gene-mapping'.

DNA sequencing of the A*2502 allele confirmed the polymorphism at nucleotide position 282, distinguishing it from A*2501 (figure 6.8). As with many other HLA alleles, the nature of sequence polymorphism found within this allele probably arose through a gene conversion or recombination event (Parham, et al 1988), since the defining polymorphisms found within A*2502 are present in other HLA-A alleles. As with A*2501, this variant contains the Bw4 polymorphism. This is found between amino acid positions 77-83, which is located at the top of the alpha helix of the alpha-1 domain of the class I molecule. Expression of the Bw4 motif has been shown to inhibit lysis by a subset of natural killer (NK) cells containing the NKB1 receptor (Gumperz, et al 1995). Although slightly different to HLA-B molecules containing the Bw4 motif, HLA-A*2403 and A*2501 have been shown to prevent lysis by the NKB1 subset of natural killer (NK) T cell clones. It therefore seems likely that A*2502 will also exhibit this NK inhibitory capability.

reaction	SSP	posn.	size (bp)	1	2	3	4
1st	#38	2.71C	691	+	+	+	+
	#AC	3.216C					
a	#4	2.188C	532	+	+	+	+
b	#18	2.197A	523	-	-	-	-
c	#6	2.209G	511	-	+	+	+
d	#34	2.226C	494	+/-	-	-	-
e	#44	2.228G	492	-	+	-	-
f	#45	2.228A	492	+	-	+	+
g	#2	2.229A	491	-	-	-	-
h	#11	2.229G	491	+	-	+	+
i	#32	2.238C	482	-	+	-	-
j	#33	2.244G	476	-	+	-	-
k	#15	2.246C	474	+	-	+	+
l	#21	2.268A	452	+	+	+	+
anchor	#AE	3.174A		+	+	+	+

Table 6.4. PCR gene-map of A*2502 (3,4) in comparison to A*2501 (1) and A*6601 (2). A*2502 is identical to A*2501, except for the polymorphisms defined by AL#6. Some crossreactivity was noted with AL#34 and A*2501.

The name A*2502 was officially assigned by the WHO Nomenclature Committee and the nucleotide sequence data submitted to the GenBank EMBL database under the accession number X97802.

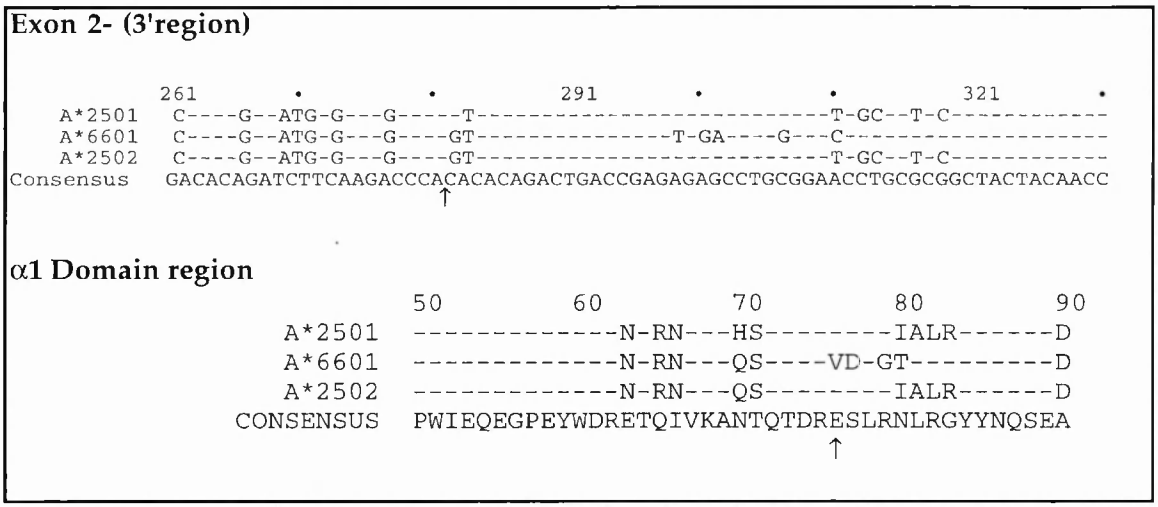


Figure 6.8. Sequence alignments showing both nucleotide and amino acid comparisons between HLA-A*2502 and A*2501, A*6601. Nucleotide sequence numbered from the first residue in exon 1, protein sequence numbered from the first amino acid in the a-1 domain. HLA-A*2502 may have resulted from a gene conversion event introducing the polymorphism found at nucleotide position 282 in A*6601 (↑) into the A*2501 sequence.

6.4. PCR-SSP typing HLA-A locus 'blanks'.

The development of PCR-SSP as applied to HLA analysis provided the opportunity to reassess the nature of HLA-A locus 'blank' specificities. Serological analysis in past workshops has revealed a significant number of HLA-A blanks (Imanishi, et al 1992a), which may reflect the inadequacy of serology for defining specificities in one population, when applied to another. This thesis has encountered two opportunities to characterise 'blank' or non-expressed HLA-A specificities. One corresponded to a new allele (HLA-A*8001) recently identified, for which serological identification proved difficult. However, this A*8001 allele (Rosen-Bronson, et al 1992; Starling, et al 1994) is expressed and can be detected serologically (Rosen-Bronson, et al 1992) with the appropriate sera. The other study contained a number of individuals all determined as possessing an HLA-A locus blank through family studies. This latter group of individuals appears to carry a

HLA-A allele which is not expressed.

6.4.1. HLA-A*8001 confirmation in a Moroccan Family.

HLA-A*8001 was identified in a number of studies (Rosen-Bronson, et al 1992; Domena, et al 1993; Starling, et al 1994). All of these studies identified this allele in black North Americans. The observed antigen frequency of this allele in the African-American population was estimated at 2% in a study of 254 individuals tested, and shown to be absent in 305 Caucasians (Starling, et al 1994). It seems likely that this specificity is of African origin which did not spread to Europe and Asia in early population migrations (Domena, et al 1993). In examination of the sequence of A*8001, it seems likely that A*8001 represents a separate sixth family of HLA-A specificities, in addition to the A3, A9, A2, A10, A19 families (Domena, et al 1993).

The lack of available sera for definition of this allele potentially can result in it being identified as a HLA-A 'blank'. Such a situation arose in serological typing of a Moroccan family (for a matched bone marrow translplant), which seemingly contained a HLA-A locus blank in one paternal haplotype (personal communication from Dr I. Joosten, Academisch Ziekenhuis Nijmegen). Serological observations identified the presence of HLA-A30 with some weak HLA-A1 reactivity. The haplotypes determined in the paternal sample were Ax;B51;Bw4 and A30;B18;Bw6, with Ax denoting the 'blank'. Dr. Joosten further typed this sample by PCR-SSP using a basic low resolution panel similar to that described in Table 3.2 in chapter 3.2.4. The observation made in this PCR-SSP typing, which did not include a reaction to A*8001, typed this sample as A*30 only. The failure to account for the weak A1 reactivity, suggested that a specificity was being missed in the PCR-SSP typing, and this may represent a novel allele.

At the time that these observations were made, the A*8001 allele was described in the black North American population. In examination of the sequence of this allele, it was clear that it represented a specificity distinct from the other A locus specificities defined in the low resolution typing panel, and would not be amplified by any of the existing SSP combinations. However sequence similarities were present with other A locus alleles, in

particular A*0101, although 34 nucleotide substitutions exist between the two alleles (Domena, et al 1993). The noted weak serological A1 crossreactivity in the Moroccan sample, made A*8001 a candidate for the observed blank specificity.

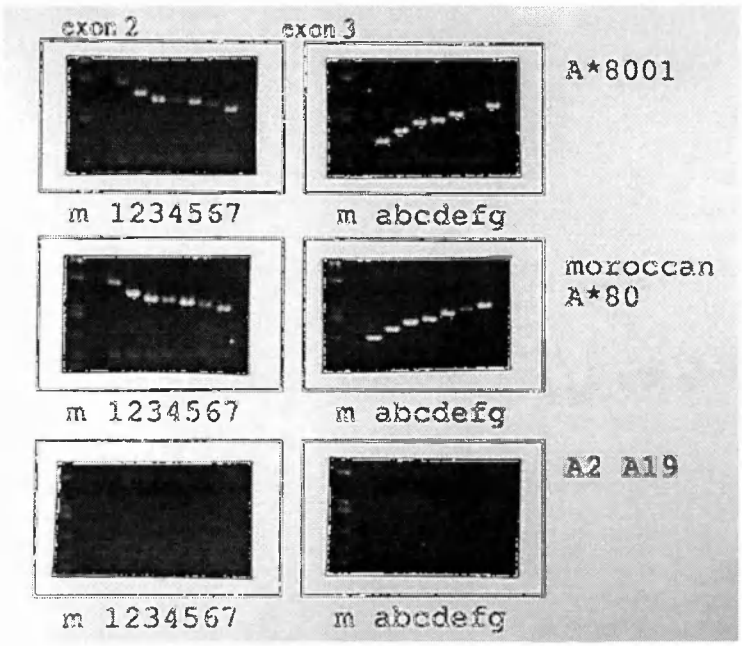
primer	motif	3' nucleotide position	A*8001	Moroccan A*8001	A2 A19
exon 2 (primers in combination with AL#Z)					
(1) AL#51	TcacA	29	+	+	-
(2) AL#19	CcA	127	+	+	-
(3) AL#8	GA	184	+	+	-
(4) AL#49	TG	199	+	+	-
(5) AL#30	acT	210	+	+	-
(6) AL#32	GggaC	238	+	+	-
(7) AL#21	ggA	268	+	+	-
AL#53	gcA	103	+	+	-
AL#54	ctA	219	+	+	-
exon 2 (primers in combination with AL#54)					
(a) AL#AB	TA	20	+	+	-
(b) AL#AM	GG	71	+	+	-
(c) AL#L	cg	110	+	+	-
(d) AL#AI	T	125	+	+	-
(e) AL#S	A	159	+	+	-
(f) AL#T	CG	184	+	+	-
(g) AL#BK	GA	217	+	+	-
AL#Z	CG	228	+	+	-

Table 6.5. Summary Table of A*8001 PCR gene mapping in the Moroccan family study. For the sequence motif column, lowercase denotes nucleotides identical to the consensus sequence, with uppercase showing polymorphism. Primer sites are shown in Figure 6.10.

With the sequence of A*8001 published (Domena, et al 1993), it was possible to design PCR-SSP combinations which would specifically amplify A*8001. The Moroccan A locus blank DNA was tested against two A*8001 specific SSP combinations (AL#53/AL#BK and AL#54/AL#BK). In addition, an A*8001 control DNA was tested, together with the other members of the Moroccan family, none of which contained the blank. Both PCR-SSP reactions identified the A blank as A*8001, which was further verified by the

A*8001 control DNA.

Since A*8001 seems to represent a novel specificity amongst HLA-A alleles, it is possible that variants could exist as amongst the other HLA-A families. To address this within the context of this Moroccan A*8001, a limited PCR 'gene-map' was performed, which identified key polymorphisms in both exon 2 and 3 of the A*8001 gene. An initial first round A locus specific PCR, amplifying the majority of exons 2 and 3 was used as template for a subsequent nested panel of 'gene-mapping' reactions identifying polymorphisms in this region of the A locus sequence. The PCR 'gene-map' as shown in figure 6.9 and Table 6.5, shows that the A*8001 characterised in this study is identical to the A*8001 control DNA at the primer sites tested (figure 6.10). It is possible that polymorphisms may exist at other locations within the sequence not considered by this panel of primers, and DNA sequencing should be considered to investigate this further.



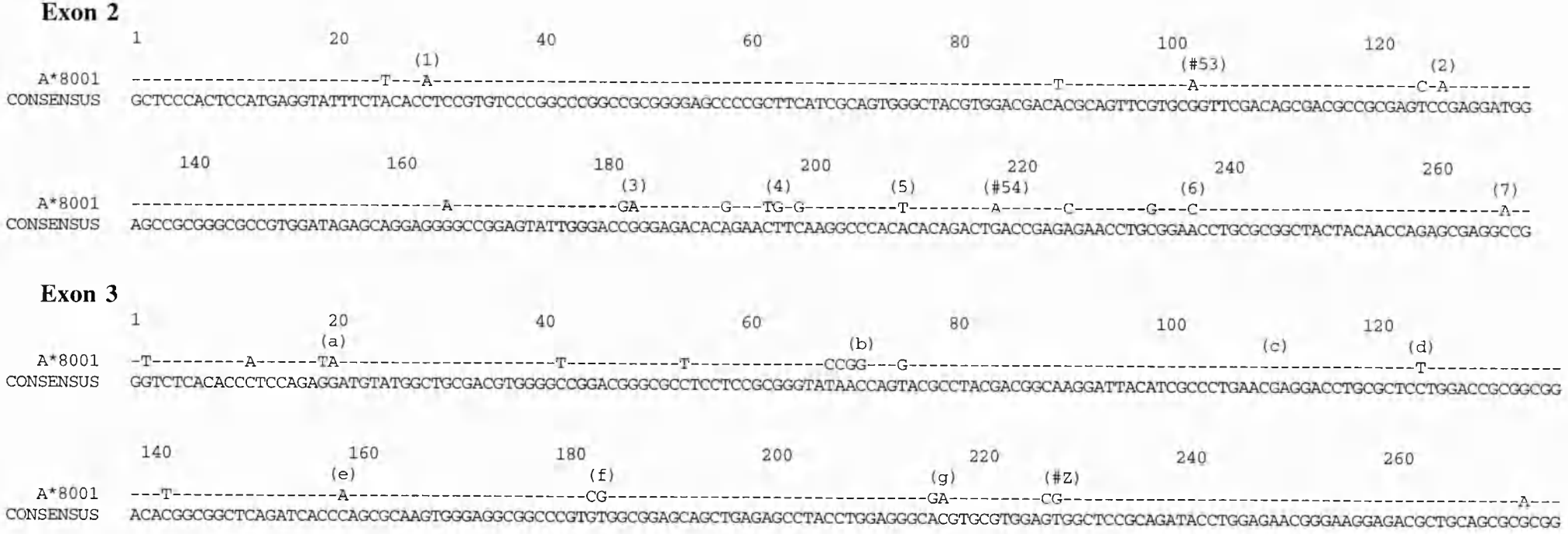
Exon 2
1. AL#51
2. AL#19
3. AL#8
4. AL#49
5. AL#30
6. AL#32
7. AL#21

Exon 3
a. AL#AB
b. AL#AM
c. AL#L
d. AL#AI
e. AL#S
f. AL#T
g. AL#BK

Exon 2 SSP in combination with AL#Z, exon 3 with AL#54

Figure 6.9. Gel showing PCR gene-map to characterise A*8001 from a sequenced A*8001 DNA, a Moroccan A*8001 individual and an A*8001 negative DNA. The results are given in Table 6.6.

ILA-A*8001 showing primer sites.



igure 6.10 The above figure shows priming sites for PCR gene mapping in exons 2 and 3 of the A*80 allele found in a Moroccan family. Sites (1-7) and (a-b) are described in ıe attached table and on the gels.

In separate studies, we have noted the presence of A*8001 at low gene frequency in typing Gambian and some Senegal individuals (data not shown). A*8001 was not identified in typing the 90 individuals in the Ugandan study (chapter 5.4, Table 5.5). Although numbers and populations studied are small, from current observations it seems likely that the distribution of this allele may vary between different parts of Africa. Future typing of presently untested populations may reveal its true distribution. Since the means for its identification now exist, future studies may resolve the presence of any additional polymorphism within the A*8001 specificity.

6.4.2. Characterisation of a HLA-A locus blank in three independent families.

During the course of this work, three independent cases of serological HLA-A locus 'blanks' have been characterised by the PCR-SSP based approach. In each case, the presence of a HLA-A locus blank was serologically determined through family studies. To achieve additional information as to the nature of the HLA-A serological blank, the use of PCR-SSP was applied to characterise these seemingly non-expressed specificities at the DNA level.

(i). The B-S family.

The B-S family was the first study in this project in which the nature of a serologically negative HLA-A specificity was investigated. The origin of the family was from Belfast. DNA from two members of the family were provided by D.Middleton, Belfast, N.Ireland. The serologically determined HLA types of the two individuals is given in Table 6.6.

	Serological Class I type							
Sample	A	A	B	B	Bw4	Bw6	Cw	Cw
JB	A3	AX	B7	B60	-	Bw6	Cw7	Cw10
JS	A3	AX	B27	B60	Bw4	Bw6	Cw1	Cw10

Table 6.6. Serological typing of two family members containing a HLA-A blank specificity.

The HLA-A locus was investigated by PCR-SSP using a low resolution typing panel as described in Table 3.4 in chapter 3. Both DNA samples typed as HLA-A3, A9. It was therefore apparent that the specificity whose

expression was not detected was HLA-A9. Further analysis by additional reactions showed that the A 'blank' specificity was typing as A*24. This was confirmed by a limited gene-mapping panel looking at four definitive sites of A24 sequence polymorphism in exon 2 and three in exon 3 (data not shown).

(ii). The A-E Family.

In trying to find a match for a bone marrow transplant patient, a Saudi family were serologically HLA typed. This exercise found a haplotype within which the HLA-A specificity that was not detected. This haplotype was present in the patient's mother and two of his siblings. Details of the family and their serological types are given in figure 6.11. The four individuals in the family possessing the serological A locus blank were typed by PCR-SSP at the A locus. The typings showed, as with the B-S individuals, the A 'blank' typing as A*24. These samples were subtyped to determine whether the A locus blank was A*2402 or A*2403. A panel of three reactions were used, which identified; A*24 (AL#8/AL#R), A*2403 (AL#8/AL#W) and A*2301, A*2402, A*2404, A*2405 and A*2406 (AL#9/AL#Z). These reactions are listed in Table 4.13. The results identified the presence of A*2402.

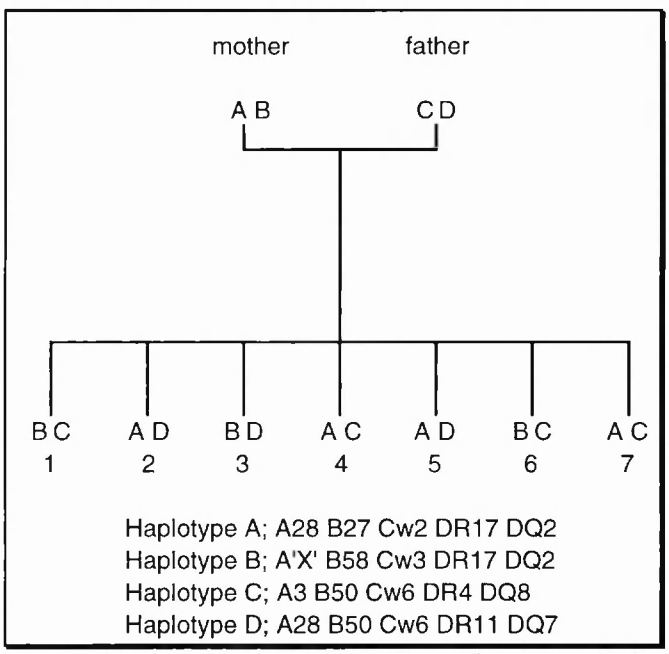


Figure 6.11. Family tree of the A-E family showing HLA haplotypes. Haplotype B contains the serological HLA-A blank (A'X'), present in individuals 1,3,6 and their mother.

(iii). The C-B family.

Three DNA samples obtained from an Israeli family had been serologically HLA typed to reveal the presence of an A locus blank. As with the other two families, these were investigated by PCR-SSP at the A locus with the blank specificity typing as HLA-A9. Further subtyping determined that the blank typed as A*2402 (data not shown).

The three above studies of serological HLA-A locus blanks found that at the DNA level, each blank could be identified as HLA-A*24. HLA-A*24 from the BS family and another HLA-A*24 blank sample (not included in this study) have recently been DNA sequenced (A*2409N, A*2411N (Marsh, et al 1996)). The exact sequence details remain as yet unpublished. However, each of the two non-expressed alleles contain a single base pair difference from the A*2402 (P.Parham, personal communication). It is interesting that in three different families, the A locus blank found was attributed to a non-expressed variant of HLA-A*2402. Also that two different non-expressed variants of A*2402 have been identified.

The loss of HLA class I expression through somatic mutation, has been associated as a possible means of tumour escape from immune surveillance (Browning, et al 1996). The explanation for non-expression of HLA-A in a germ-line, and if this incurs any susceptibility to cancer, remains to be deduced.

Other HLA-A null alleles have also been described including an A*0201 allele which failed to be expressed through a mutation in the untranslated regulatory region 5′ region of the gene (Balas, et al 1994). A*0215N is not expressed through a mutation causing a stop codon and truncation of the class I heavy chain (Ishikawa, et al 1996). Another class I allele, A*0303N fails to be expressed through a six base pair deletion in exon 3. This deletion removes a structurally important cysteine from the mature heavy chain which prevents the proper folding and so functioning of the class I molecule (Lienert, et al 1996).

The reasons for the existence of HLA-A 'null' alleles may also suggest that under certain circumstances there may be some selective advantage of loss of expression of a particular specificity. The exact relevance may be

understood better once its frequency, distribution and any disease associations with the null alleles are fully characterised.

6.5. Discussion.

The application of PCR-SSP to tissue typing has provided a novel perspective for the determination of HLA specificities, which has resulted in new alleles being identified. As a consequence of the work of this thesis, several new alleles have been characterised, their identification relying on the recognition of novel patterns of reaction in the PCR-SSP panel, or a discepancy in result in comparison with another method of typing. As the definition of polymorphism at the DNA level becomes more and more established, so an increasing number of new alleles will be found.

The characterisation of HLA polymorphism has also been addressed through the development of PCR 'gene-mapping'. PCR gene mapping offers a means of rapidly screening new variants to predict their sequence polymorphism. It is not an alternative to sequencing, which is still necessary for characterising new alleles. The data obtained from 'PCR gene mapping' in this thesis has enabled the design of more than one specific SSP combination for identification of a new variant.

Characterisation of the composition of sequence polymorphism by PCR 'gene-mapping' has been possible because of the nature of polymorphism in the HLA class I genes. The generation of polymorphism in the HLA genes seems to have arisen through the mechanisms of gene conversion or double recombination rather than point mutation, in the examples identified in this thesis. The implication of this is seen in the shared nature of polymorphic sequence motifs between HLA specificities. The unique nature of an allele is the result of the combination of these different polymorphic sequence motifs along the length of the gene. This suggests that there is an apparent finite number of polymorphic sequence motifs found amongst the HLA class I alleles. Defining the HLA specificity on the basis of these sequence motifs, provides a rapid and up to now, accurate approach to predicting the DNA sequence of alleles, as demonstrated in the use of PCR-SSP to tissue type, charactersise or confirm new specificities.

This chapter also described the use of PCR-SSP to define serologically HLA 'blank' specificities. In one study this was simply due to a novel HLA-A allele (A*8001), which was serologically ill-defined. With the sequenced published, it was possible to define the presence of A*8001 by PCR-SSP and also confirm its sequence composition by PCR 'gene-mapping'. Another study involving identification of a serological blank in three independent families, produced an interesting observation. By PCR-SSP, the serological blank in all three families, typed as A*2402. Mutations have since been reported which account for the non-expression of the A*24 gene sequenced (A*2409N, A*2411N (Marsh, et al 1996)).

The ability to identify the non-expressed A*24 specificity by PCR-SSP does cause some concern. Typing these individuals by PCR-SSP alone would not draw attention to the fact that it was not expressed. This raises an argument that tissue typing in the context of the expressed HLA molecule (i.e. serology or IEF), does hold relevance in the future of histocompatibility testing. The exact frequency of such cases, which appear rare may detract from this argument. However, as seen in this chapter, the comparison of observations made by different approaches to HLA typing does provide an objective view from which novel specificities can be identified.

The development and application of new typing approaches have therefore expanded the knowledge and perception we have of the HLA system. As more sequence based approaches are applied to identification of HLA polymorphism in different populations, so it seems plausible that the identification of an increasing number of new alleles will follow. This will undoubtedly determine the true extent of HLA polymorphism, which will greatly benefit the fields of anthropology, disease association and transplantation.

Chapter 7

Determining the significance of polymorphism amongst the A*30 allelic group.

7.1. Introduction.

The purpose of this thesis has been to establish a method for tissue typing at the HLA-A locus, which allows definition to an allelic level. This has been demonstrated in the preceding chapters through the ability of SSP primers designed on the ARMS principle to distinguish between alleles of highly similar sequence in the PCR-SSP reaction. Having established the ability for such discrimination, it was also important to contemplate the relevance of typing to such a high level of resolution.

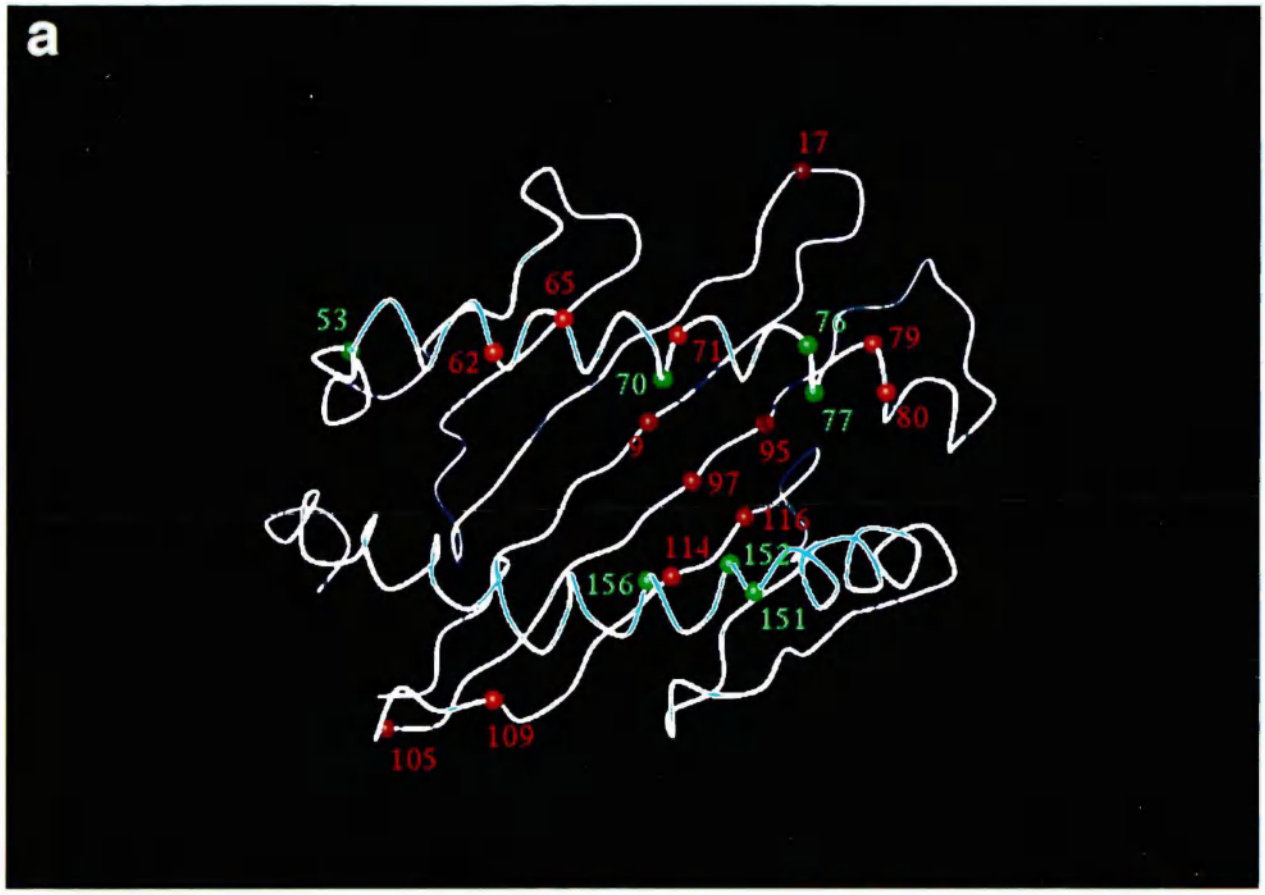
In comparing the DNA sequences of the class I alleles (Arnett and Parham 1995), it is clear that the majority of polymorphism lies within exons 2 and 3. The significance of this becomes apparent when the location of these polymorphisms are considered within the expressed class I molecule. Exon 2 and exon 3 code for the alpha-1 and alpha-2 domains of the class I molecule. These two domains form a binding groove which presents peptide to the TCR. Changes within this peptide binding groove have potential relevance as to the sequence of peptide bound together with its orientation. This in turn may elicit a different HLA restricted T cell response.

To explore the importance of sequence polymorphism in terms of the presentation of peptide by the class I molecule, a group of alleles, serologically defined as HLA-A*30 were compared. HLA-A*30 alleles have been extensively considered in other parts of this thesis; their identification

through nested PCR in chapter 4.4.3., their distribution in the Sardinian and Ugandan populations in chapter 5, and the identification and characterisation of a new allelic variant, A*3004, in chapter 6.3.1.

HLA-A*30 was identified as a split of a serologically cross-reactive group of antigens known as HLA-A19. HLA-A*29, A*31, A*32, A*33 and A*74 constitute the other members of this group (Kato, et al 1989). HLA-A*30 has been included as part of the A19 group, mainly on the basis of its serological cross-reactivity with A*31. The serological definition of A*31 in the presence of A*30 has proved almost impossible, due to the specificity of available sera being limited to A*30 specific or A*30/A*31 specific, with no mono-specific sera for A*31 being easily available (Carpenter, et al 1989). The serological definition of A*31 in the presence of A*30 has therefore relied on family studies. The basis of this serological A*30 cross-reactivity with A*31 lies with the sharing of an Arg_{56}, which lies at the top and at the end of the alpha-1 α-helix (figure 7.1a). This site would seem to be antigenic and accessible to binding of antibody.

Early molecular analysis of the A19 group through DNA sequencing (Kato, et al 1989), established that HLA-A30 (A*3001) was probably part of the A*01/A*03/A*11 family of specificities, rather than the A19 group. Inclusion of A*30 in the A19 serological group (as discussed above), was by virtue of its cross-reactivity with A*31. The isolation of A*30 from the other members of the A19 group was further corroborated through early observations made in the development of the PCR-SSP system at HLA-A locus. In endeavoring to design a single PCR reaction to encompass the A19 group of specificities, it proved impossible to find an SSP combination that would also include A*30. The low resolution typing panel as described in chapter 3 (Table 3.4), therefore determined the A19 serological group through two reactions. One identifying A19 with the exception of A*30, and the other amplifying only A*30. The A19 not A*30 amplification is achieved through the specificity of SSP AL#F. This primer identifies a silent polymorphism at nucleotide position 105 in exon 3. This polymorphism is found uniquely in the A19 not A*30 allele group, and further substantiates the exclusion of A*30 from this group of specificities.


a
17
65
53
62
71
76
79
70
77
80
9
95
97
116
114
152
156
151
109
105

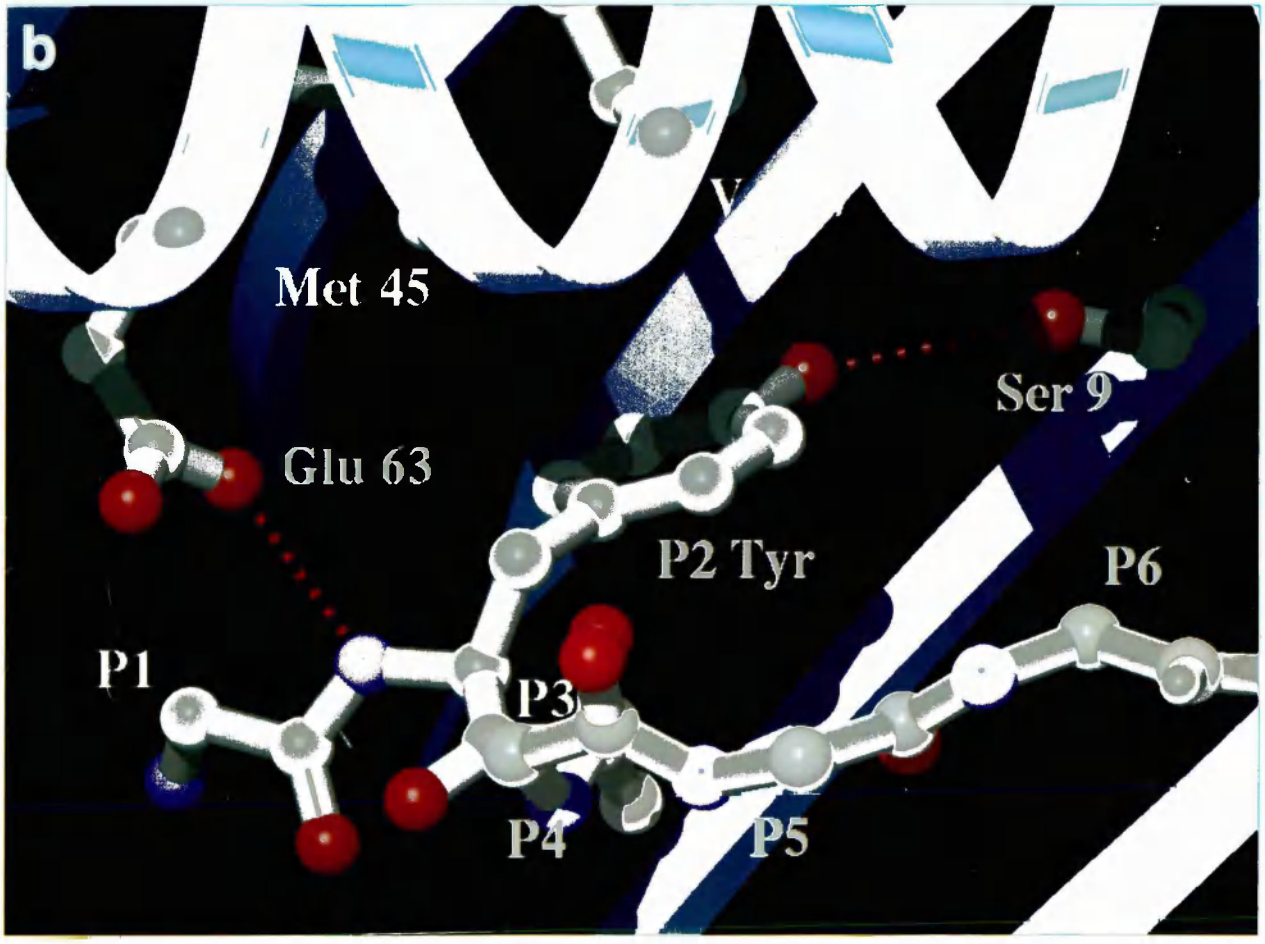

b
Met 45
Glu 63
Ser 9
P2 Tyr
P6
P1
P3
P4
P5

Figure 7.1. (as shown on previous page). (a). A ribbon diagram showing positions of polymorphism within the four A*30 variant molecules. Positions are marked as per the mature protein. Positions marked in red denote polymorphic differences shared amongst the A*30 group as compared to a consensus sequence (Arnett and Parham 1995). Positions marked in green denote sites of polymorphic differences between the A*30 variants. Note erratum; position 53 is marked incorrectly, this polymorphism should be at position 56, located at the top of the loop directly above position 53. Details of the polymorphism marked in green are given in Table 7.3. (b). A representation of peptide binding into the HLA-A*3002 molecule within the B pocket. The peptide sequence (AYSSWMYSY) was reported to bind and be presented by A*3002 (Alan Rickinson, personal communication). The detail shows the side-chain of the peptide P2 Tyr binding into the B pocket, forming a hydrogen bond with the Ser at position 9 of the A*3002 heavy chain. Both figures were generated in MOLSCRIPT (Kraulis, 1991, modification by R.Esnouf), and rendered with Raster3D (Merritt, et al 1994).

Early IEF analysis of HLA-A*30 revealed three variants (Yang, 1989a). Of the three IEF variants, A30.1 and A30.3 were found to be the most common, with A30.2 rarely seen. A fourth even rarer variant, A*3003, is also distinguishable by IEF (Choo, et al 1993). Figure 6.3 in chapter 6 provides examples of all four IEF definable A*30 variants. Analysis showed that of the two common variants, A30.1 is associated with B*18, while A30.3 is associated with B*13 in Caucasoids, and B*42 in Black populations. These two main IEF variants have now been DNA sequenced, as have the two rarer variants (Bodmer, J.G. et al 1995).

When considering cellular approaches to vaccine strategies, accurate tissue-typing is essential to establish the prevalent HLA restricted T cell mediated immune responses. With the view of a potential future HIV vaccination study, a cohort of 90 HIV positive individuals from a Ugandan semi-urban population, were extensively HLA typed by PCR-SSP (chapter 5.4.). In this population, HLA-A*30 represented the most common allelic group amongst the HLA-A locus specificities. Unlike HLA-A*02 (Rotzschke, et al 1992; Barouch, et al 1995; Sudo, et al 1995), very little is known with regard to peptide binding and presentation by the different allelic variants of HLA-A*30. Knowing the respective peptide motifs of the different allelic variants of HLA-A*30 is therefore an important part of any cellular based vaccine studies in which the target population contains a significant frequency of A*30 individuals.

This study therefore considered the relevance of the polymorphism found amongst the four variants of HLA-A*30 in terms of peptide binding. To

ascertain the functional differences between the four A*30 allelic variants, the peptide binding preference of each allele was determined. This required the construction of HLA transfectants, in which the different A*30 variants were expressed in a HLA class I negative cell background. Endogenously processed peptides were then eluted from the different expressed A*30 alleles and the preferred peptide sequence determined. The similarities and differences observed in the peptide motifs eluted from the different A*30 variants, provide additional insight into the relevance of HLA polymorphism found within the peptide binding groove. As discussed, this also provides useful information for future cellular based vaccine studies.

7.2. Construction of the HLA-A*30 transfectants.

7.2.1. Cloning and transfection.

The basic procedure for making transfectants for the four variants of HLA-A*30 is as described in chapter 2. The BCL as listed in Table 7.1, represent cell lines in which the four allelic subtypes of HLA-A*30 have been identified through DNA sequencing. These BCL provided the A*30 genes which were transfected and expressed in the class I reduced (C1R) cell line. In brief, this involved first preparing mRNA from growing cells, and then making first strand cDNA through RT-PCR. The full length HLA-A cDNA was amplified using specific primers as detailed in figure 2.2 in chapter 2. This HLA-A PCR was then ligated into a pMOSblue T vector to facilitate cloning of the A*30 alleles. The ligated vector was then used to transform competent cells and clones were picked, grown and plasmid DNA prepared. This DNA was then screened by PCR-SSP using an A*30 specific primer combination (AL#12/AL#G). DNA amplifying with this SSP pair were identified as A*30. These clones were then DNA sequenced. Only one clone contained a number of errors from the published sequence. Further sequencing of additional clones confirmed the published sequence and suggested incorporation of Taq errors in the problematic clone.

Having confirmed the published sequence for all four subtypes, the HLA insert was cut out of the cloning vector and ligated into the pKG4 expression vector. This required the use of an intermediary 'shuttle vector'

(pBluescript), due to a lack of compatible enzyme restriction sites between the cloning and expression vectors. At each of these steps, clones were picked and plasmid DNA subtyped for A*30 by PCR-SSP (see chapter 4.4.3). HLA typing the plasmid DNA at each stage confirmed both the presence and subtype of the A*30 insert (figure 7.2).

BCL	HLA-A		HLA-B		HLA-C	A30 IEF subtype	Accession Number
LBF	A*3001		B*1302		Cw6	A30.3	M30576
CR-B	A*3002	A*6602	B17	B18		A30.1	X61702
JS	A*3003	A3	B7	B18		-	M93657
W7(CC)	A*3004	A11	B49	B37		A30.2	Z34921

Table 7.1. HLA types of the B cell lines from which the A*30 allelic variants were cloned. Corresponding A30 IEF subtypes are also given. EMBL accession numbers for each of the A*30 sequences are also given.

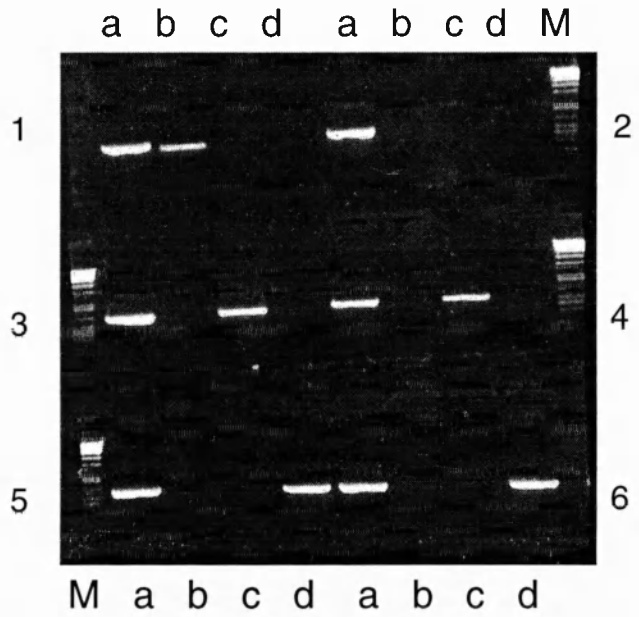


Figure 7.2. PCR-SSP typing screen to verify A*30 clones. Reaction a=A*30 (AL#12/AL#G), b=A*3001 (AL#6/AL#K) c=A*3002 (AL#10/AL#T) d=A*3003 (AL#10/AL#U), with SSP combinations in brackets. Lane M denotes the size marker. All bands are in the region of 350bp, consistent with the expected size for amplification of cDNA. Hence, clone 1 types as A*3001 (positive with a+b), clone 2 types as A*3004 (positive with a only), clones 3 and 4 type as A*3002 (positive with a+c) and clones 5 and 6 type as A*3003 (positive with a+d).

Having established the presence of all four A*30 variants in the pKG4 transfection vector, a maxiprep preparation of the plasmid DNA was made and purified through a CsCl gradient. This purified plasmid DNA was subsequently linearised (using restriction enzyme ScaI) and electroporated into growing C1R cells. C1R cells have reduced surface expression of HLA-A and B molecules (Edwards, et al 1982; Storkus, et al 1987; Zemmour and Parham 1992). The plasmid DNA was also digested with the HindIII, BamH1 restriction enzymes used to ligate into the vector and viewed on an agarose gel (figure 7.3). This was used as a simple screen to ensure the integrity of the insert and vector prior to transfection. Cells were then cultured in selection and transfectants observed growing out after a period of two to three weeks following electroporation.

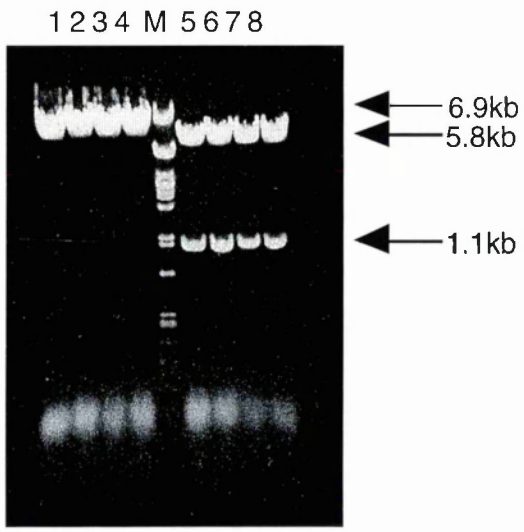


Figure 7.3. Restriction digests to verify integrity of the A*30/pKG4 expression vector constructs prior to transfection. Digestion is shown for the A*3001 construct (lanes 1 and 5), A*3002 (lanes 2 and 6), A*3003 (lanes 3 and 7), A*3004 (lanes 4 and 8). Lane M is the size marker. Verification of linearisation of constructs with restriction enzyme ScaI (lanes 1-4) is seen as a single 6.9kb band. Digestion with HindIII/BamH1 restriction enzymes (lanes 5-8) excises the A*30 inserts (1.1kb) from the pKG4 vector (5.8kb). HindIII and BamHI restriction sites facilitated ligation of the A*30 cloned cDNA into the pKG4 expression vector.

7.2.2. Verification of expression.

Expression of the transfected A*30 was verified through FACS analysis. Unfortunately, no A*30 specific antibody was available for this screening and the HLA class I specific monoclonal W6/32 (Brodsky, et al 1979), was used for used to detect HLA class I expression (figure 7.4). More specific screening of expression was possible in conjunction with Mike Bunce at the Oxford Transplant Centre. The four transfectants were analysed for expression by serology using the complement mediated lymphocytotoxicity

assay (Terasaki and McClelland 1964; Bodmer and Bodmer 1979), as described in chapter 1.2.1.

The serological results (Table 7.2) of this characterisation provided a perception of the relationship between the detection of HLA-A30 by antibody and sequence polymorphism. Screening with the reaction panel produced two predominant patterns of reactivity. Firstly, using antibodies FE82 and MP6, lysis was observed with the A*3002, A*3003 and A*3004 but not the A*3001 transfectants. The specificity of this sera and monoclonal antibodies had been described as A*30 positive but negative for the IEF variant A30.3 (Mazzoleni, et al 1991). As discussed above, the A30.3 IEF variant corresponds to A*3001. The results observed with the four A*30 transfectants and the A*3002 (IEF A30.1) B cell line confirmed the described specificity of FE82 and MP6. In comparing the amino acid sequence of all four A*30 subtypes, a number of possible polymorphic differences exist which may explain these observations. A*3001 differs from the other A*30 variants by three amino acid substitutions in the alpha-1 domain. These substitutions are $His_{70} \rightarrow Gln_{70}$, $Glu_{76} \rightarrow Val_{76}$ and $Asn_{77} \rightarrow Asp_{77}$ in A*3001, as listed in Table 7.3. These substitutions are found within the α-helix of the alpha-1 domain, and are at positions potentially accessible to antibody (figure 7.1a).

The other serological pattern of reactivity observed was positive with A*3001, A*3002 and A*3004, but negative with A*3003. This pattern of reactivity was present with some of the sera used in Table 7.2. As discussed in the introduction to this chapter, sera detecting A*30 often cross-reacted with A*31. In viewing the amino acid sequences of A*30 and A*31, this can be explained in part by the sharing of a common polymorphism (Arg_{56}) found in the alpha-1 domain. Interestingly, at this position A*3003 differs from the other A*30 subtypes and A*31 by having a Gly. The negative result obtained with A*3003 with several of the sera therefore maps the specificity of the A*30/A*31 sera to this polymorphism in the alpha-1 domain. This site is on the α-helix at the end of the peptide binding groove, and should be accessible to antibody (figure 7.1a).

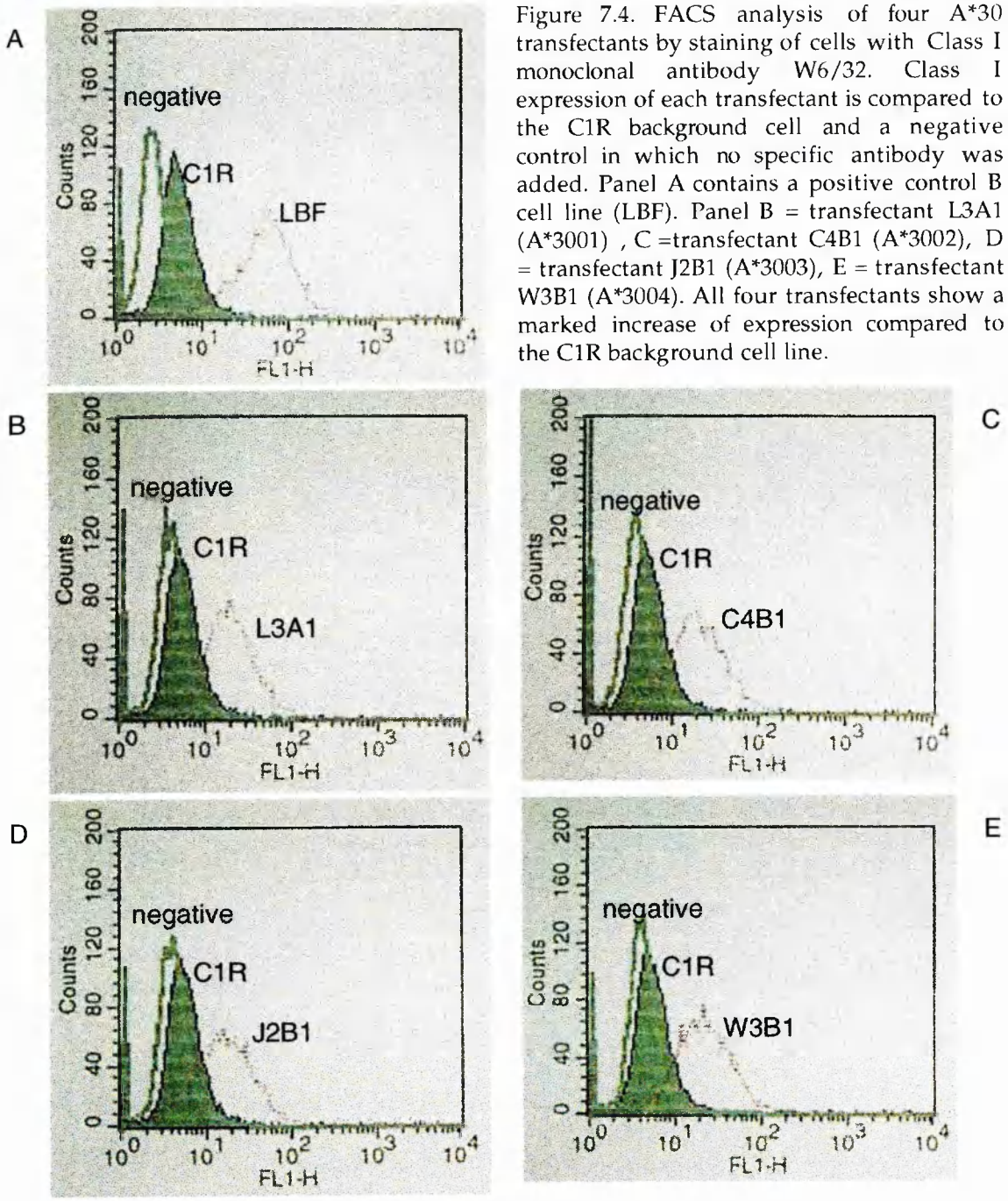


Figure 7.4. FACS analysis of four A*30 transfectants by staining of cells with Class I monoclonal antibody W6/32. Class I expression of each transfectant is compared to the C1R background cell and a negative control in which no specific antibody was added. Panel A contains a positive control B cell line (LBF). Panel B = transfectant L3A1 (A*3001) , C =transfectant C4B1 (A*3002), D = transfectant J2B1 (A*3003), E = transfectant W3B1 (A*3004). All four transfectants show a marked increase of expression compared to the C1R background cell line.

A third pattern of reactivity had been reported whereby A*3002 (IEF A30.1) was negative while the other A*30 alleles were positive (Carpenter, et al 1989). This pattern of reactivity was not observed with the panel of sera used in this characterisation. The above results do however indicate that the transfectants are expressed, and conformationally correct within the context of the A*30 sera specificity. One anomaly was observed with antisera 14948,

being positive with the A*3002 BCL (CR-B), but negative with the corresponding transfectant. This observation may be due to the lower expression of A*30 on the transfectant, or a different than reported specificity for the sera or an error in typing.

		Transfectants				Controls	
Sera	Specificity	L	C	J	W	C1R	CR-B
FE82	A30,Cw5 (IEF A30.3 neg)	0	4	3	3	0	4
MP6*	A30, (IEF A30.3 neg)	0	4	4	4	0	4
Myl102	A30	4	4	0	3	1	4
S15488	A30, A31, A33, A29	4	4	2	4	0	4
Kinloch	A30, A31, A33, A29	3	4	2	2	0	4
8788	A30, A31, A33, A29	3	4	0	3	0	4
14948	A30, A31, A33, A29	0	0	1	0	1	4
Tome	A30 A31	4	4	3	4	2	4
1:1	" "	4	4	1	4	0	4
1:3	" "	4	4	0	3	0	4
1:7	" "	3	3	0	3	0	3

Table 7.2. Serological characterisation of the four A*30 transfectants (L=A*3001, C=A*3002, J=A*3003, W=A*3004). The transfectants, together with a negative (C1R) and A*30 positive cell line (CR-B=A*3002) were tested against a panel of A*30 antisera and monoclonal antibody (*) in a complement mediated lymphocytotoxicity assay. The sera Tome was tested in the given dilutions to avoid crossreactivity as seen when used neat. Reactivity with the sera was registered using a scoring system (0-4). A score of 0 = no significant lysis above background, 1 = around 10-25% lysis above background, 2 = around 25-50% lysis above background, 3 = around 50-75% lysis above background and 4 = 75-100% lysis above background.

	alpha-1 domain				alpha-2 domain		
	56	70	76	77	151	152	156
A*3001	R	Q	V	D	R	W	L
A*3002	R	H	E	N	R	R	L
A*3003	G	H	E	N	R	R	L
A*3004	R	H	E	N	H	V	W

Table 7.3. A comparison of the amino acid differences between four A30 variants in the alpha-1 and alpha-2 domains of the HLA Class I molecule.

7.2.3. Peptide elution and sequencing.

The process of peptide elution was performed in collaboration with Professor Ramensee's laboratory at the Institut Für Zellbiologie, Eberhard-Karls-Universität, Tübingen, Germany. The trifluoroacetic acid (TFA) elution method was performed essentially as described previously (Falk, et al 1991) and briefly outlined in chapter 2.3.6.

A comparison of the peptides eluted from each of the four HLA-A*30 alleles, provided an opportunity to consider the influence of the defining polymorphic differences separating the four variants. It also allowed the comparison of HLA-A*30 as a group, in comparison to other HLA-A specificities.

HPLC analysis of the eluate from the HLA class I column identified several peaks which were absent in the non-specific glycine control. These peaks were observed between 15 and 50 minutes following loading of the sample onto the HPLC. Edman sequencing of the material contained within these peaks identified several individual peptides endogenously processed by the C1R transfectants (Table 7.4). Additionally, the fractions (not including the largest peaks and potentially non-specific fractions) were pooled for sequencing.

7.3 Results of peptide elution.

The results of Edman degradation sequencing of the peptides eluted from the four A*30 variants are shown in Table 7.4. Sequencing of the pooled peptide fractions showed that peptide motifs existed (Falk, et al 1991) for each of the four A*30 molecules. It was also possible to identify individual ligands for each of the four A*30 variants (Table 7.4). Two ligands were identified for A*3001, four identified for A*3002, four for A*3003 and five for A*3004. The probable proteins from which these peptides may have originated are also listed. One ligand identified in A*3003 (HBBF) was derived from a bovine protein. This peptide represents the endogenous

A*3001	P1	P2	P3	P4	P5	P6	P7	P8	P9					
anchors		**F**							**F**					
preferred		LY		P	D		V		N					
ligand	Y	**F**	D	P	A	N	G	K	**F**					EF2 (265-73)
	K	T	K	D	I	V	N	G	L					CAPB (235-243)

A*3002	P1	P2	P3	P4	P5	P6	P7	P8	P9					
anchors		**VL Y**							**Y**					
preferred		MF	H K	PE D	I			K						
ligands	A	**Y**	K	K	Q	F	S	Q	**Y**					RL5 (217-25)
	R	I	S	G	V	D	R	Y	**Y**					unknown
	R	**L**	A	H	E	V	G	W	K	**Y**				R13A (139-48)
	R	**V**	L	P	P	N	W	K	**Y**					RS30 (132-40)

A*3003	P1	P2	P3	P4	P5	P6	P7	P8	P9					
anchors		**F**							**Y**					
preferred	GI	ILQ VW Y	N	P	DE RH									
ligands	R	**I**	S	G	V	D	R	Y	**Y**					unknown
	S	**L**	F	V	S	N	H	A	**Y**					ALFA (355-363)
	S	G	H	A	I	I	A	Y	**Y**					unknown
	V	**V**	T	G	V	A	N	A	L	A	H	R	Y	HBBF Bovine (132-44)

A*3004	P1	P2	P3	P4	P5	P6	P7	P8	P9					
anchors		**QT VY M**							**Y**					
preferred		FI L	A N KF	PD G	E	IL M	HY		MR VF	**Y**				
ligands	A	**Y**	K	K	Q	F	S	Q	**Y**					RL5 (217-25)
	K	**M**	K	E	I	A	E	A	**Y**					HS71 (126-34)
	K	**Y**	F	D	E	H	Y	E	**Y**					CKS2 (11-19)
	H	**T**	S	D	V	P	G	K	L	I	**Y**			ITB2 (520-30)
	K	**Q**	F	A	E	M	Y	V	A	K	**F**			EF2 (226-36)

Table 7.4. Legend on following page.

Table 7.4. (shown on previous page) Results of Edman sequencing of peptides eluted from the four A*30 transfectants. P1 to P9 represent residue positions within the peptide. Both individual sequences of ligands and anchor and preferred residues determined from pool sequencing are shown. Amino acid residues depicted in bold represent proposed anchor residues within the peptide sequence.

processing of protein found in foetal calf serum introduced through the tissue culture media. This phenomena still represents normal endogenous peptide processing and has been observed on other occasions (Christian Münz, personal communication).

The data detailed in Table 7.4 identifies anchor preferences for all four A*30 variants at positions P2 and P9 (amino to carboxyl) within the peptide sequence. Individual peptides longer than 9 amino acids (between 10 and 13 residues) were sequenced, with the terminal amino acid corresponding to the deduced anchor residue. Anchor residues were assigned on the basis of a greater than 50% increase and decrease of signal in the previous and following cycles respectively (raw data not shown). The peptide anchor residues for A*3001 identify a strong preference for amino acids F at both P2 and P9. Preferred residues of L and Y are also noted at the P2 position. A*3002, A*3003 and A*3004 differ from A*3001 by having a preference for Y at P9. At P2, A*3002 favours large or moderate sized hydrophobic residues (V,L,Y,M,F) which is similar to that observed with A*3003. A*3004 also prefers larger hydrophobic residues at P2 (V,Y,M) but can also accommodate Q and T as anchors.

The similarities in the determined motifs, particularly between A*3002, A*3003 and A*3004 are strengthened when considering the individual ligands identified for each of the variants. The same ligand was eluted and sequenced from both A*3002 and A*3003 (RISGVDRYY). In addition, an identical ligand was determined for both A*3002 and A*3004 (AYKKQFSQY). Both these ligands contain hydrophobic residues at P2 and the A*3002, A*3003, A*3004 shared Y anchor preference at P9.

There were preferred residues at other positions within the determined motifs for all four A*30 variants. At P3, there was an apparent preference for charged residues (H,K,R,D,N), at P4 there was some preference for smaller

residues (P,G,D) and at P5, there was a suggested preference for charged residues (D,E,). A*3004 also shows some preference for larger hydrophobic residues at P6 (I,L,M) and polar hydrophobic residues at P7 (H,Y). The preference of the A*30 molecules for these minor positions within the peptide sequence may further determine the nature of peptide bound.

7.4 Discussion.

The MHC restricted immune response consists of three main components, namely the MHC molecule, its peptide ligand and the specificity of the T cell receptor. MHC molecules exhibit specificity for the peptides they bind, with a preference for certain anchor residues within the peptide sequence (Falk et al. 1990; Falk et al. 1991). This preference for a particular motif within the peptide is related to the polymorphism in the binding site of the HLA molecule. In addition to a selectivity for peptide motifs, polymorphism within the peptide binding groove may influence the orientation of the peptide and so the antigenic surface presented to the TCR. The specificity of the TCR will relate to the orientation of a particular sequence within the context of a given MHC molecule.

HLA polymorphism. This study examined the HLA molecule peptide interaction in a group of closely related alleles. These variants of HLA-A*30 differ from each other in and around the peptide binding groove as illustrated in figure 7.1a. As summarised in Table 7.5 below, the four A*30 variants differ from each other by one to six amino acid residues. The functional relevance of these polymorphisms depends upon their position within the HLA molecule and how this influences the binding of ligand and interaction with the TCR.

	A*3001	A*3002	A*3003	A*3004
A*3001	0	4	5	6
A*3002		0	1	3
A*3003			0	4
A*3004				0

Table 7.5. A comparison of the number of amino acid differences between each of the four A*30 variants.

The polymorphic differences as depicted in figure 7.1a and listed in Table 7.3, all lie around the peptide binding groove. Several of the polymorphism found amongst the A*30 variants as a group are in positions sensitive to binding of anchor residues within the peptide sequence. Polymorphism at position 9 within the mature class I heavy chain has been shown to be important amongst the A*02 allelic group, with relation to the character of the B pocket and its interaction with binding of the P2 side chain of the peptide (Bjorkman et al. 1987; Madden et al. 1993). In the A*30 molecules, Ser occupies this position. The small size of the Ser side chain has a steric effect, creating additional space within the B pocket, enabling it to accommodate and bind bulkier side chains. Other polymorphisms at amino acid positions Glu_{114} and His_{116} play a potential role in the D and F pockets respectively. Glu_{114} may confer a preference for a P3 charged side chain residue binding into the D pocket, which is supported by the sequencing data from the eluted peptides (Table 7.4). The polymorphism and amino acid composition around the F pocket (amongst which Glu or Val_{76}, Asp or Asn_{77}, Gly_{79}, Thr_{80}, His_{116} contribute), suggest a pocket that is both large and fairly neutral.

Polymorphism between the A*30 variants, as shown in figure 7.1a and detailed in Table 7.3, may also be relevant with respect to the peptide binding characteristics. Differences between A*3001 and the other A*30 variants are present at a number of amino acid positions, including 76 and 77 (Table 7.3). This polymorphism may contribute to the P9 anchor residue differences observed between A*3001 and the other A*30 variants (Table 7.4). On the other α–helix, polymorphic differences between the A*30 variants at position 151, 152 and 156 may contribute to conformation changes in the middle of the peptide. There are three patterns of polymorphism in this region of the molecule, (as seen in (a). A*3001, (b). A*3002/3 and (c). A*3004), which may cause the peptide to bind in three different orientations within the binding groove. A*3002 and A*3003 have practically identical sequence except for a Arg→Gly change at position 56. This position may not effect the peptide binding motif, but it is solvent accessible and may have some interaction with the TCR. This position is discussed below in relation to the TCR.

The peptide motif. The Edman sequencing of the peptides eluted from the A*30 transfectants reveal a preference by all four variants for anchor residues at P2 and P9 (or terminal residue if greater than 9 amino acids in length) within the peptide. At both of these anchor residues, and in all four variants, there is a preference for binding large to moderate sized hydrophobic residues. This peptide motif fits into the observations made above as to the large B and F pockets and their accommodation of bulkier peptide side chains. The similarities of the peptide motifs can be considered in the context of the shared polymorphism which characterises the A*30 variants a group.

The P2 anchors (Phe, Val, Leu, Tyr, Gln, Thr, Met) show a preference by all four A*30 variants, for binding large hydrophobic residues into the B pocket. As mentioned, the Ser_9 in the A*30 heavy chain is common to all four A*30 variants and may facilitate this interaction by contributing to a large B pocket. HLA-A*23 and A*24 specificities also have a Ser_9. The similarity between A*30 and A*24 in the context of Ser_9 is reflected in the Tyr P2 anchor motif determined for A24 (Rammensee et al. 1995). A representation of the possible interaction between a slightly polar P2 Tyr peptide side chain and Ser_9 of the A*30 molecule, is depicted in figure 7.1b. Figure 7.1b also shows the importance of Glu_{63} in interacting with the N-terminal peptide residues.

The characterisation of the variant A*30 P9 peptide anchors revealed variation amongst the group of specificities. A*3001 shows a preference for Phe at P9 while the three other A*30 specificities all show a preference for a Tyr P9 peptide anchor. Both of these anchor residues are large, aromatic and hydrophobic, but there is a clear difference in the preferences seen amongst the A*30 group. As discussed, the dichotomy of preferences observed at the P9 anchor residue, coincides with the polymorphic differences between A*3001 and the other A*30 variants, particularly at positions 70, 76 and 77 in the alpha-1 domain (Table 7.3 and figure 7.1a).

The similarities in the sequences determined for ligands of the A*30 group, suggests that on the basis of peptide, there is scope for a TCR to crossreact between the four variants. The extent of this crossreaction will be dependent

on a particular TCR's sensitivity to HLA polymorphism and any conformational changes in peptide orientation.

The TCR. The combination of peptide bound to the class I HLA molecule provides a picture within a frame that an appropriate TCR can recognise. Within this study, it has not been possible to examine the specificity of the TCR to the different peptide-A*30 complexes. But the results obtained in terms of the peptide binding motifs together with the polymorphism that exists between the A*30 variants, suggests a possible spectrum of interactions. It is therefore interesting to speculate as to the different possible responses that may be observed between combinations of peptide, HLA-A*30 variants and HLA restricted CTL clones.

Recently, a picture of the interaction between the TCR, peptide and HLA-A2 has been determined through the crystal structure of this complex (Garboczi et al. 1996; Garcia et al. 1996). This has provided a both clear and informative insight into the relationship between all three components of the MHC restricted immune response. The following few paragraphs describe, the structure of the TCR and its interaction of peptide bound to HLA, as reported in Garboczi's paper (Garboczi et al. 1996). The results from this chapter are then considered in the light of these TCR observations.

The membrane bound TCR comprises of an alpha and beta chain each consisting of constant and variable domains (Davis and Bjorkman 1988). The variable domains are generated through a rearrangement of gene segments which allow the wide diversity observed in the T cell repertoire. Each of the variable domains consists of a number of loops or complimentarity determining regions (CDR1, 2, 3). The CDR loops of the alpha and beta chain combine to form a flat surface which interacts with the MHC molecule and bound peptide.

In determining the crystal structure of TCR-Peptide-HLA complex, the positioning of the TCR was found to be in a diagonal orientation over the HLA-A2 and peptide, with the CDR loops in line with the beta pleated sheets of the class I molecule. This configuration fits within the two 'high points' of the HLA class I molecule, located at N-terminal of the Alpha-1 domain α-helix and the N-terminal of the alpha-2 domain α-helix (Garboczi et al.

1996). The positioning of the CDR loops has also been determined in relation to the peptide and the HLA molecule. Both alpha and beta variable domains of the TCR contact both peptide and HLA class I molecule.

In general, it was found that the TCR covers a substantial area of the peptide/MHC complex and appears to push the peptide down into the binding groove. This pressing down on the peptide by the TCR causes a small but noticeable shift in the MHC conformation. The TCR covers the length of the peptide, with the TCR pocket (formed mainly by the CDR3α and CDR3β loops), aligned over the middle peptide residues.

CDR3β, the potentially most diverse of the variable loops, spans across the binding groove, associating with both helices of the class I molecule (at amino acid residues 72 in the alpha 1 domain α-helix and 149, 150, 151, 155 in the alpha 2 domain α-helix) and residues P5-P8 of the HTLV-1 Tax peptide bound in this study (Garboczi et al. 1996). The TCR central pocket covers the middle of the peptide, contacting residues P4 and P5. In addition the CDR3α loop contacts amino acid residues 65, 66, 68 and 69 of the alpha-1 domain α–helix. CDR1α contacts the amino end of the peptide and is positioned over both α-helices (at around amino acid residue 58 in the alpha-1 domain α-helix and 163, 167, 170 in the alpha-2 domain α-helix). CDR1β associates with the P8 residue of the peptide. CDR2α sits over the alpha-2 domain α-helix making contact with residues 155, 158 and 166.

The CDR2β loop is positioned over the alpha-1 domain, but makes no apparent contact with peptide or the HLA molecule. The lack of contact between CDR2β and the HLA-peptide complex may be due to the absence of a carbohydrate moiety attached to the HLA molecule in the crystallised structure. This is probably a consequence of generating the complex in bacteria. This carbohydrate is usually attached to a conserved Asn residue at position 86 in the alpha-1 domain α-helix (Barber, et al 1996), and may be important in interacting with the TCR CDR2β directly or assisting in the interaction with the alpha-1 domain of the HLA molecule (Parham, 1996). It has also been speculated that the failure of the CDR2β loop to contact a region of the HLA molecule (residues 73-84), associated with binding to receptors found on natural killer cells, may be linked to a function of these cells (Parham 1996).

Applying the observations made in the HLA-A2-Tax peptide-TCR complex to the A*30 group, provides possible insights concerning the likely interactions between the different components of the complex. HLA polymorphisms that are common across the A*30 group and are likely to contact the TCR, are found at amino-acid residues 62 and 65 in the alpha-1 domain α-helix, which as discussed above, are in a region which may contact the CDR1α and CDR3α loops of the TCR. As stated, the common A*30 polymorphisms identified at amino acid residues 79 and 80 in the alpha-1 domain α-helix do not appear to bind to the CDR2β loop, although they are proximal to each other. As discussed, this may be a consequence of the lack of N linked glycosolation to residue 86 of the class I heavy chain, or some function of NK activity (Parham 1996).

In considering the crystal structure of the TCR-Peptide-HLA interaction (Garboczi et al. 1996), certain polymorphic differences between the HLA-A*30 variants may be definitive with regard to interacting with the TCR. In viewing figure 7.1a and Table 7.3, the polymorphism that exist at amino acid residues 151 and 152, and a possible interaction with the CDR3β loop may be a key point of discernment in which the TCR exhibits its specificity. This is an important area amongst the A*30 sequences, since three patterns of polymorphism exist (Table 7.3), distinguishing A*3001 from both A*3002/3 and A*3004. The polymorphic differences at these positions within the HLA sequences may be significant in respect of defining restriction for the TCR.

As discussed, the only difference between A*3002 and A*3003 lies at amino acid position 56 in the HLA sequence. This is located in the N-terminal region of the alpha-1 domain α-helix and under Garboczi's observations (Garboczi et al. 1996), would have no direct interaction with the TCR. It is likely therefore that a TCR specific for one of these two variants would crossreact with the other. However, this is only speculation and whether a polymorphism at amino acid position 56 holds any relevance as to peptide binding or contact with the CDR1α loop of the TCR, remains to be established.

How a particular TCR and peptide interact in different A*30 variant backgrounds, will be dependent on the relevance of changes in peptide orientation within the binding groove and how sensitive the TCR is to those changes. The orientation of the peptide will in part determine which residues are accessible to the TCR. In the case of the A*30 ligands, residues which are important in contacting the TCR are found within the middle of the peptide, with P2 and P9 (or terminal residue) being employed as anchor sites for binding to the A*30 molecules. The P1 side chain may also be accessible to the TCR, since this peptide residue is non-specifically anchored into the A pocket of the HLA class I molecule through the amino group on its main chain. As discussed, the TCR covers the majority of the peptide and potentially can be sensitive to subtle changes in peptide conformation. In this respect, it seems plausible that a TCR specific for a particular peptide bound to say A*3002 or A*3003, could be sensitive to changes in the peptide's orientation if bound to A*3001 (through the influences of polymorphism at amino acid residues 70, 77, 152) or A*3004 (through the influence of polymorphism at residues 152 and 156 in the class I heavy chain).

It is also feasible that a TCR specific for peptide bound within one HLA-A*30 variant, will recognise the same peptide bound to one or more of the A*30 variants. This is possible because of the similarities in peptide motifs determined for all four A*30 variants. In addition, a TCR may be tolerant of polymorphic differences between each of the A*30 alleles, which would manifest itself as crossreactivity by the TCR in recognising more than one of the A*30 variants.

In summary, this study has compared the preferences for peptide binding of four variants of HLA-A*30. The sequencing of peptide from each of the four A*30 variants has characterised both similarities and differences in the preferred ligands. All four variants of A*30 have preferred anchor residues at P2 and P9 (or terminal carboxyl residue if longer). All four A*30 variants preferred binding peptides with large to moderate sized anchor residues. However, a clear difference was noted at the P9 (or terminal residue if longer), with A*3001 preferring a Phe anchor while the other A*30 variants preferring a Tyr. Although the A*3002, A*3003, A*3004 have similar peptide motif preferences, the role of the polymorphic differences that define them may have significant influence on peptide orientation and MHC restriction.

It was not possible to address these functional issues with regard to TCR specificity within this thesis, but this will hopefully be achieved as part of an ongoing study.

This study both demonstrates and contemplates the relevance of polymorphism with regard to the functioning of the HLA molecule in the MHC restricted T cell response. The study also shows that very similar HLA molecules have different peptide motif preferences, and speculates that as with variants of A*02 (Rotzschke, et al 1992; Barouch, et al 1995; Sudo, et al 1995), T cell responses will be both specific and cross reactive, depending on the requirements of the TCR. The differences observed in peptide binding and the suggested effects of HLA polymorphic differences on peptide orientation and presentation to the TCR amongst this group of A*30 molecules, supports the requirement to type at a resolution that identifies the functionally distinct HLA molecule.

Chapter 8

HLA-A*02 polymorphism and function.

This chapter is based on a review article written by the thesis author for the European Journal Immunogenetics, 1996: 23: 261-274.

8.1. Introduction.

HLA class I molecules bind and present peptides derived from endogenously processed proteins. These peptides are presented to $CD8^+$ T cells, with the interaction restricted through HLA type. The extent of polymorphism and so the variety of HLA specificities, is a striking characteristic of the HLA system. Serological discrimination of HLA types has demonstrated this variety through the specificity of antisera. But serology by no means revealed the true level of polymorphism, which became increasingly apparent through the fine discrimination of the T cell in HLA restricted immune responses (Gotch, et al 1985; Latron, et al 1991; Rotzschke, et al 1992; Utz, et al 1992, van der Poel, et al 1986; Tussey, et al 1994; Barouch, et al 1995a; Sudo, et al 1995). DNA sequencing and developments in DNA-based HLA typing have shown polymorphism well beyond that defined through serology, as witnessed through recent nomenclature reports (Bodmer, J.G. et al 1995). The majority of polymorphism detected at the DNA level is important in the context of causing amino acids changes in regions of the HLA molecule, sensitive to the discriminatory binding and presentation of peptide.

A good demonstration of the effect of serologically silent polymorphism in relation to the function of the HLA molecule can be found amongst the

allelic variants of HLA-A*02. HLA-A*02 has been well defined serologically and is the most common HLA type in many population groups. But this one serological specificity has been shown to consist of a number of allelic variants. At the time of writing, some sixteen expressed variants of HLA-A*02 have been described (Bodmer, 1995), (a seventeenth, A*0215N is not expressed). In addition, the distribution of HLA-A*02 variants differs, according to the population group studied (Carcassi, et al 1995; Krausa, et al 1995). Many of the polymorphisms which distinguish the HLA-A*02 variants have been shown to have functional relevance in terms of binding and presentation of peptide to the T cell receptor (Gotch, et al 1985; McMichael, et al 1988; Barouch, et al 1995a; Sudo, et al 1995).

This chapter will describe the nature and possible evolution of polymorphism found amongst the HLA-A*02 alleles. It will further discuss how the distribution of HLA-A*02 subtypes varies in different population groups. The significance of HLA-A*02 polymorphisms with regard to peptide binding and the T cell mediated immune response, will also be considered. The sequence polymorphism observed between the A*02 alleles can be considered as a possible response to pathogenic pressure which gives further insight into the diversity of the HLA system.

8.2. The nature of HLA-A*02 polymorphism.

Seventeen (A*0201-A*0217) allelic variants of HLA-A*02 have been described (Bodmer, 1995) at the time of writing. Most have been DNA sequenced, (A*0208 A*0209 have been amino acid sequenced). In comparing the sequences, it is clear that the majority of polymorphism is found around the peptide binding groove of the HLA molecule (figure 8.1, figure 8.2), and so has potential relevance in determining the sequence characteristics and conformation of the bound peptide. However, A*0209 and A*0215N differ from other A*02 alleles through polymorphic differences in exon 4, both probably as a result of point mutation. A*0215N is not expressed due to this polymorphism leading to a stop codon and so truncation of the heavy chain midway through the alpha-3 domain.

HLA-A*02 alleles contain polymorphisms which clearly define them as a group of related alleles and explain their serological specificity. In the alpha-

1 domain, there are a number of polymorphisms restricted to the A*02 allelic group, and in the alpha-2 domain, there are several sites of polymorphism shared between A*02 and A28 alleles (figure 8.1). The presence of such polymorphic sites explains serological observations in the clear identification of the HLA-A*02 specificity and its crossreactivity with A28 alleles. In addition to this, there exists an array of polymorphisms which define the differences between HLA-A*02 allelic variants, few of which are discriminated serologically. As illustrated in figure 8.1, the majority of polymorphic differences between A*02 allelic gene products cluster around the peptide binding groove. Consequently, many of the polymorphic differences between A*02 alleles have potential functional significance in peptide presentation.

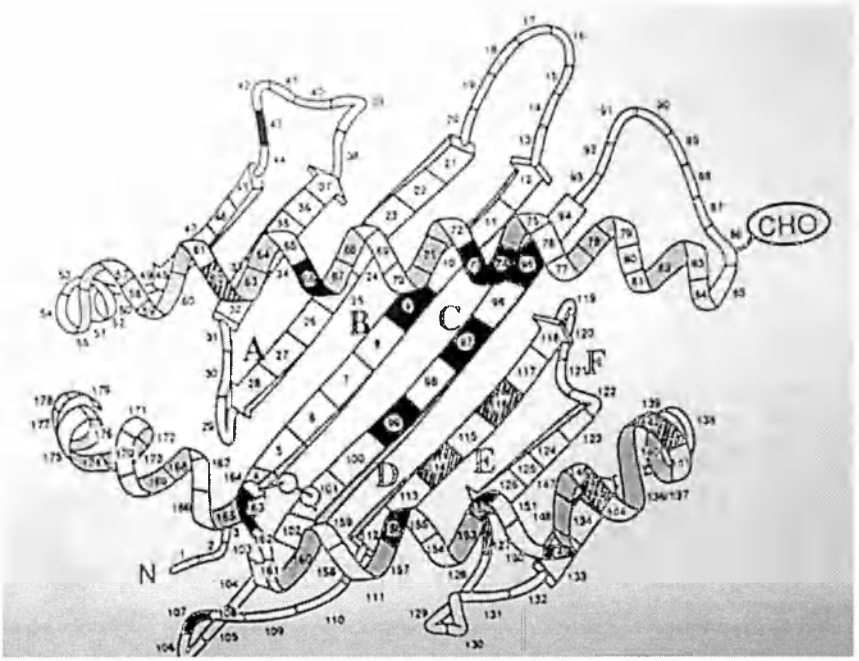


Figure 8.1. A ribbon diagram representing the α-1 and α-2 domains of the class I molecule, which together form the peptide binding groove. Location of pockets within the groove is shown (A-F). Polymorphic amino acid positions which differentiate between the different A*02 alleles are represented as solid boxes. Striped positions identify sites of polymorphism which define the HLA-A2/28 broad specificity. The polymorphism at amino acid position 62 is specific to HLA-A*02 at the A locus, positions 142, 145 are specific for A2/A28 and positions 114, 116 and 127 are specific for A2/A28/A9.

The nature of polymorphism which distinguishes the majority of HLA-A*02 alleles, seems to have arisen through the process of gene conversion, or recombination events (Parham, et al 1988; Erlich, et al 1991b; Geraghty, et

al 1992; Kuhner, et al 1992). Hence polymorphism observed in a particular hypervariable region is often shared with other alleles, usually but not necessarily from the same locus. As examples, the polymorphism observed in A*0216 at amino acid position Glu_{163} is also found in A*6602, A*8001, B*0702-5, B*1301-2, B*2702-9, B*4001-4,6-8, B*7301, B*8101, Cw*0201-2, Cw*1701. It therefore appears that A*0216 has arisen as a gene conversion event with one of these alleles (Barouch, et al 1995b). Also, the polymorphism observed in A*0217 at amino acid position Leu_{95}, Met_{97}, Phe_{99} is also found amongst A*24 alleles. A*0217 was found in a South American Warao population, in which A*24 is present at high frequency. This suggests A*0217 might have resulted as a gene conversion event between an A*24 allele and an A*02 allele (Selvakumar, et al 1995). In viewing alignments of class I sequences (Arnett and Parham, 1995), it is apparent that gene conversion is a probable major generator of HLA diversity not only between A*02 alleles but for class I as a whole (Wells and Parham, 1996).

A comparison of the nucleotide sequences of A*02 alleles suggests they were all derived from a common ancestral allele which may or may not be one of the variants already identified. Alleles which have certain polymorphic elements related to HLA-A*02 have been described in the gorilla Major Histocompatibility Complex (Lawlor, et al 1991). Thus, the polymorphism found within the HLA-A*02 alleles may have been generated over a long period. In examining the possible relationship between HLA-A*02 alleles as a function of probable gene conversion events or point mutations (figure 8.3), it is clear that two groups of alleles exist. One group clusters around A*0201, the other around A*0205. The distinction between the two groups is apparent, with the A*0205 group (A*0202, A*0205, A*0208, A*0214) differing from A*0201 at amino acid positions $Gln_{43} \rightarrow Arg_{43}$ and $Val_{95} \rightarrow Leu_{95}$ (figure 8.2). In particular, the polymorphism at Arg_{43} is unique to the A*0205 group. The relation between the A*0205 group is further strengthened, through the presence of a silent polymorphism, probably the result of a point mutation, at nucleotide position 59'G' in exon 3 (consensus 'C'). The silent polymorphism implies that the A*0205 allelic group arose from a common ancestral allele which subsequently diversified through processes of gene conversion or recombination.

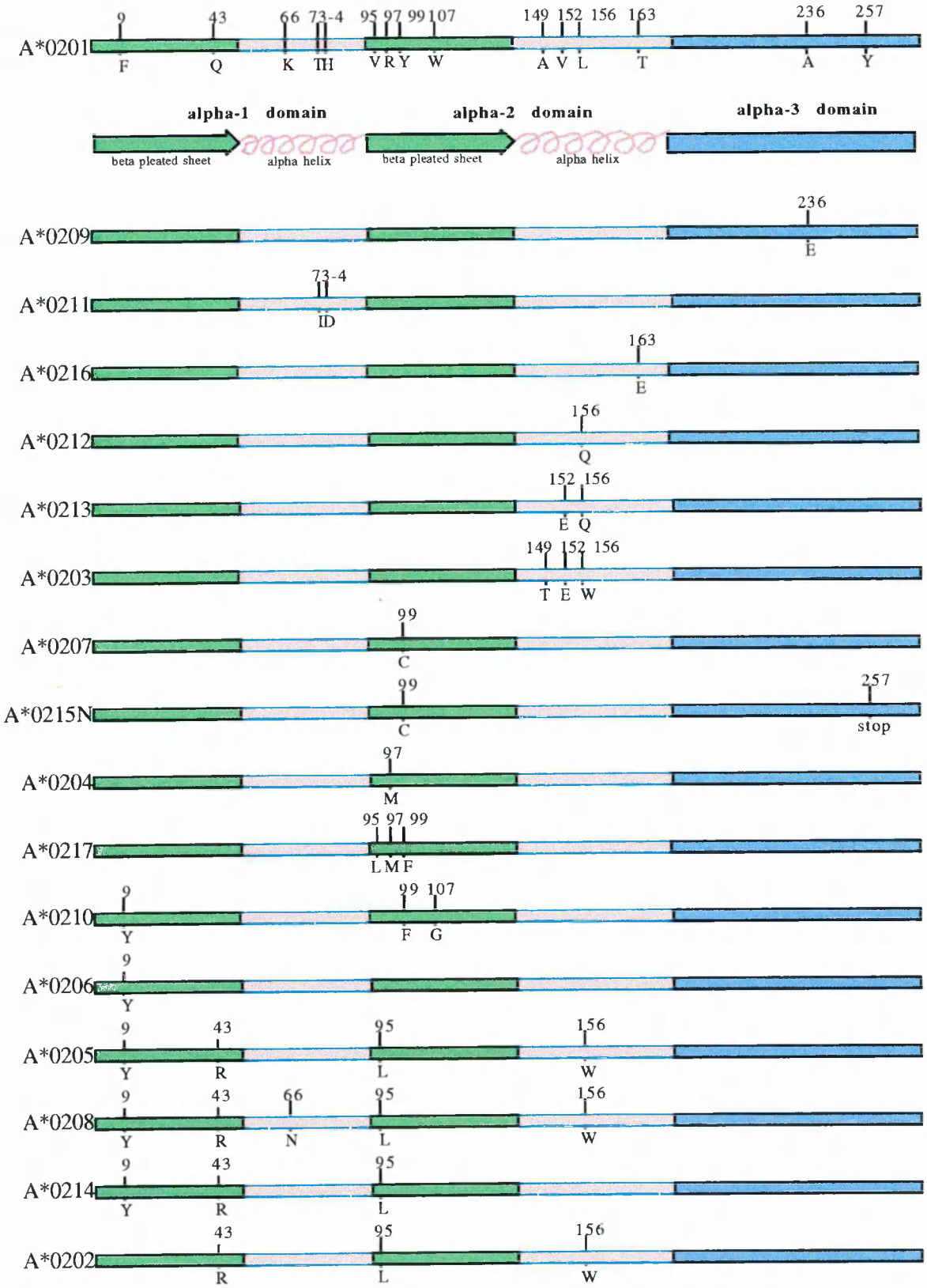


Figure 8.2. A comparison of A*02 allelic amino acid sequences. The polymorphisms given for each A*02 allele are differences observed from A*0201. The location of the polymorphism is given in the context of its amino acid position, the domain and whether it is found in the alpha helices or beta pleated sheet of the peptide binding groove.

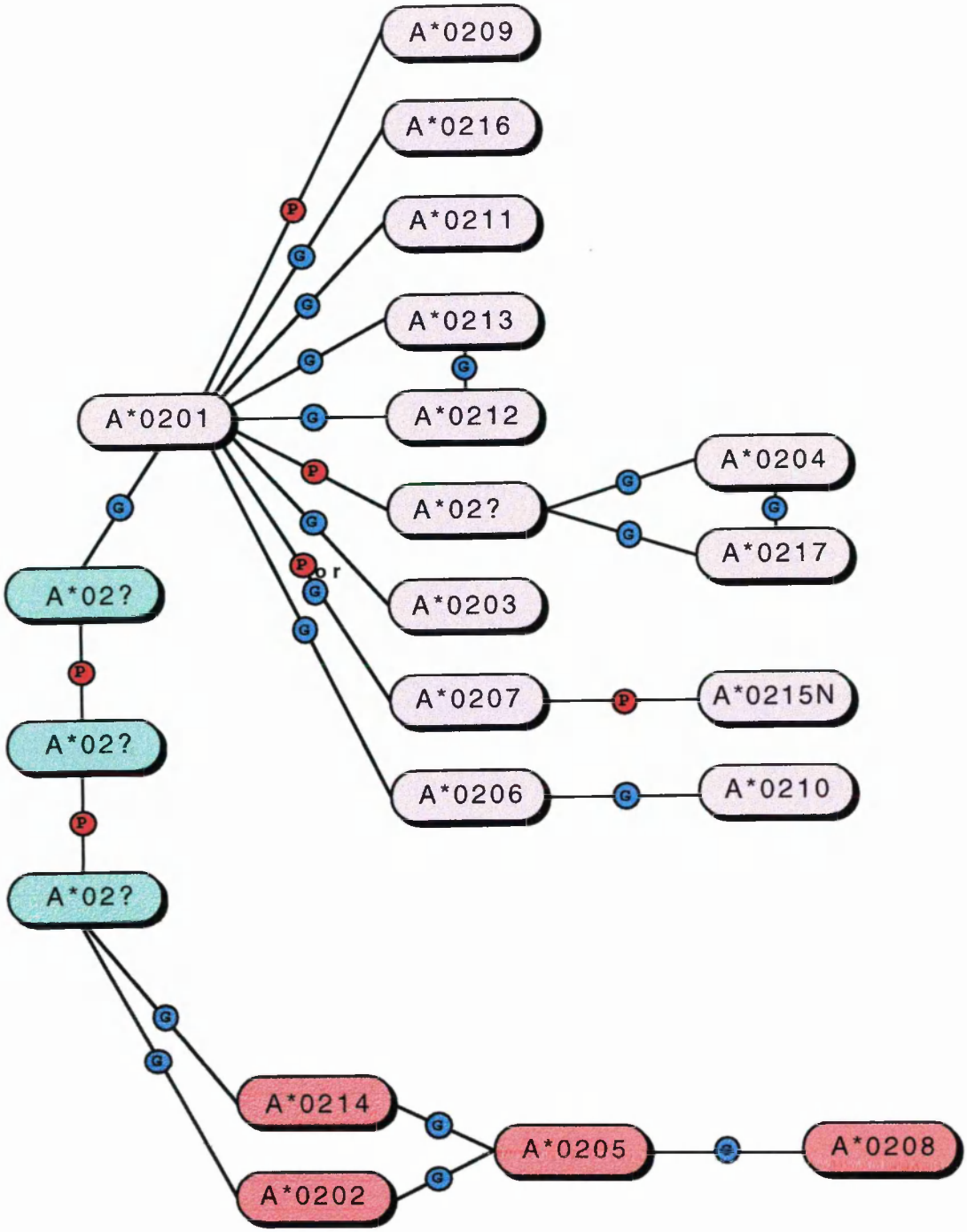


Figure 8.3. A representation of the possible relationship between A*02 alleles in terms of likely gene conversion events (G) or point mutations (P). It can be seen that the alleles split into two groups, one that can be derived from A*0201, the other which clusters around A*0205. The two groups are separated by a number of unknown intermediate alleles (A*02?), which may have been lost through selective pressures or remain unidentified.

Figure 8.3 also suggests the existance of undiscovered intermediate allelic variants which would explain those HLA-A*02 alleles currently found. These intermediate forms may have been lost through lack of selective pressure, or simply not yet identified. It seems probable that, as increasing numbers of individuals and ethnic groups are investigated by high resolution typing and molecular approaches, new HLA-A*02 variants will be detected. This will possibly lead to clarification of the overall evolutionary relationship between the HLA-A*02 alleles.

8.3. HLA-A*02 allelic variation in different populations.

In examining polymorphism within the context of ethnic origin, it is clear that very different HLA-A*02 allelic frequency structures exist between different populations. Chapter 5.3 describes these difference in reference to a number of key population groups, which are summarised in Table 5.2. The prominence of A*0201 and relative lack of A*02 allelic diversity in the Oxford Caucasian population is in contrast with the diversity observed in the Singapore Chinese population. Not only is A*0201 replaced by A*0207 as the most prominent A*02 allele in the Singapore HLA-A*02's tested, but several alleles (A*0203, A*0206, A*0210) are not seen in the other population groups tested (Krausa, 1995). Similarly A*0202 and A*0214 are only observed in the African populations typed, and A*0204 and A*0217 appear specific to American Indian populations (Selvakumar, 1995).

Even between Caucasian populations, difference in frequencies of HLA-A*02 alleles exist. This is especially reflected in a comparison between the Oxford Caucasian and Sardinian populations (Carcassi, et al 1995; Krausa, et al 1995) as shown in Table 5.2. The high proportion of A*0205 in the Sardinian population as compared to the Oxford population underlines the potential HLA heterogeneity present between different population groups of similar ethnicity. Also, within the different African populations, different HLA-A*02 frequency structures are observed, with A*0214 seemingly restricted to Uganda and Kenya, and not seen in the Gambia. These observations are limited however, in view of numbers tested.

The general extent of this allelic variation is far from known, since it is only recently that the ability to discriminate definitively between such variants has been possible. However, the observed heterogeneity of HLA-A*02 allelic frequencies already suggests a difference in pathogenic pressures driving HLA polymorphism within the different population groups. Such a mechanism has been proposed for acquiring protective immunity in a HLA restricted cytotoxic T cell response against Malaria in an African population (Hill, et al 1992; Aidoo, et al 1995). The A*02 allelic heterogeneity observed in the Sardinian population may in part be explained by the past history of endemic malaria on the island (Piazza, et al 1985). Therefore, the prevalence of HLA alleles within different populations may provide an insight as to role of such alleles in terms of the T cell mediated immune response against a particular localised pathogenic pressure.

8.4. HLA-A*02 structure.

The structural definition of the HLA-A2.1 molecule was elucidated by X-ray crystallography (Bjorkman, et al 1987). It revealed an HLA molecule which is capable of presenting a short peptide bound within a groove formed by two 'parallel' alpha helices sitting on a base of beta pleated sheets. This HLA class I molecule is composed of an HLA-encoded heavy chain glycoprotein with transmembrane and cytoplasmic region, associated with beta-2 microglobulin light chain, with the bound peptide completing the trimeric structure. The peptide binding domains of the molecule is formed by the alpha-1 and alpha-2 domains coded for by exons 2 and 3 of the HLA gene. The alpha-1 and alpha-2 domains sit on an alpha-3 domain further attached to a transmembrane region and cytoplasmic tail. The alpha-3 domain complexes non-covalently with beta-2 microglobulin, forming a platform which supports the alpha-1 and alpha-2 domain peptide binding groove.

Additional structural resolution of HLA-A*02 revealed six pockets (A-F) within the groove (Saper, et al 1991); see figure 8.1. These six pockets have potential to bind amino acid side chains present in the peptide sequence. Table 8.1 shows some of the important amino acid residues with regard to these pockets. The properties of the amino acid residues which form the pockets within the binding groove will determine the nature of the anchor residue it may bind. Hence, a pocket may be deep or shallow, contain

charged, polar, neutral or hydrophobic residues, all of which will be reflected in the compatibility of the anchor residue in relation to its side chain characteristics (Matsumura, et al 1992).

In comparing the expressed A*02 alleles in Table 8.1, it becomes apparent that, with the exception of A*0209, all the A*02 alleles listed contain polymorphisms in one or more pockets which render them distinct. These differences will have functional implications as to peptide bound and its conformation. In comparing A*02 as a group, polymorphism also exists within all the pockets as compared to the class I consensus sequence. This again will have a broad effect on the A*02 specificity of peptide binding. Only the F pocket does not contain polymorphism between HLA-A*02 molecules, which probably explains the conservation of the dominant C-terminal peptide anchor residue as Leu or Val for all A*02 molecules tested. Additionally, polymorphism at amino acid positions 107 and 149 have been implicated in binding with the T cell receptor (Saper, et al 1991). The comparison of peptide characteristics against HLA polymorphism provides an insight into the mechanics of HLA-ligand interactions.

8.5. Functional significance of variation between HLA-A*02 alleles.

The discrimination of the T cell provided evidence of the existence of functionally distinct subtypes of HLA-A*02 (Gotch, et al 1985). Alloreactive CTL specific for A*0201 failed to kill another A2 cell, later shown to be A*0205. The determination of the A2 crystal structure made apparent the importance of the polymorphic differences between these A2 subtypes in relation to the peptide binding groove (Bjorkman, et al 1987; Saper, et al 1991).

8.5.1. The influence of the peptide.

Exploring the mechanisms for binding of peptide to HLA class I provided an increased understanding of HLA resticted CTL reaction. The specificity of peptide binding was shown to be a function of the six pockets (A-F) in the peptide binding groove (Saper, et al 1991; Matsumura, et al 1992; Colbert, et al 1993; Fruci, et al 1993; Madden, et al 1993; Kubo, et al 1994; Tanigaki, et al

1994; Tussey, et al 1994). It was shown that peptide was bound in a particular conformation, with the A pocket binding the amino group and the F pocket binding the Carboxyl end of the peptide (Latron, et al 1991). The ends of the groove were also shown to be closed, which restricted the length of bound peptide, usually to a nonamer or decamer. This also has the effect of causing the peptide to bulge out of the groove, making residues more accessible to interaction with the T cell receptor (Saper, et al 1991; Matsumura, et al 1992).

Characterisation of peptides eluted from class I molecules showed that particular amino acids were preferred at certain positions within the peptide sequence (Falk, et al 1991; Rammensee, et al 1995). These preferred positions were termed anchor residues and shown to associate with one or more of the pockets (namely B,C,D,E and sometimes F) within the peptide binding groove. Only a few residues within a particular peptide sequence constitute anchor residues, so only a few pockets within the groove tend to be involved in specific binding of peptide anchor residues. This implies that other pockets can accommodate a large number of different peptide residues, increasing the number of different peptides a particular HLA molecule can bind (Matsumura, et al 1992). The prevalence of specific anchor residues within the general protein pool will also determine the number of possible peptides a particular HLA molecule may bind. If a particular HLA molecule's preferred anchor residues are rarely found amongst protein sequences, then this will limit the number of peptides that HLA molecule can bind and present (Kast, et al 1994). In the case of HLA-A*0201, its anchor residues are commonly found and so a large array of peptides can be presented (Falk, et al 1991). This may reflect the high frequency of HLA-A*0201 present across a large number of different population groups.

The sequence of the bound peptide is an important component of the interaction between the HLA class I complex and the T cell. It was shown that altering the sequence of a A*0201 specific flu matrix peptide at certain positions could have a number of different effects (Gotch, et al 1988). Certain changes had no influence on CTL recognition. Other changes in peptide sequence caused it not to be bound. However certain peptide analogues were shown to bind, yet abrogate recognition by the CTL clone. These observations therefore suggested that certain peptide residues were

	A pocket							B pocket							C pocket					D-E pocket				F pocket		
A2	59	62	63	66	99	163	167	9	45	63	66	67	70	99	9	73	74	95	97	99	114	152	156	77	80	116
A*0201	Y	G	E	K	Y	T	W	F	M	E	K	V	H	Y	F	T	H	V	R	Y	H	V	L	D	T	Y
A*0202	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	L	-	-	-	-	W	-	-	-
A*0203	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	E	W	-	-	-
A*0204	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	M	-	-	-	-	-	-	-
A*0205	-	-	-	-	-	-	-	Y	-	-	-	-	-	-	Y	-	-	L	-	-	-	-	W	-	-	-
A*0206	-	-	-	-	-	-	-	Y	-	-	-	-	-	-	Y	-	-	-	-	-	-	-	-	-	-	-
A*0207	-	-	-	-	C	-	-	-	-	-	-	-	-	C	-	-	-	-	-	C	-	-	-	-	-	-
A*0208	-	-	-	N	-	-	-	Y	-	-	N	-	-	-	Y	-	-	L	-	-	-	-	W	-	-	-
A*0209	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
A*0210	-	-	-	-	F	-	-	Y	-	-	-	-	-	F	Y	-	-	-	-	F	-	-	-	-	-	-
A*0211	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	I	D	-	-	-	-	-	-	-	-	-
A*0212	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Q	-	-	-
A*0213	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	E	Q	-	-	-
A*0214	-	-	-	-	-	-	-	Y	-	-	-	-	-	-	Y	-	-	L	-	-	-	-	-	-	-	-
A*0216	-	-	-	-	-	E	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
A*0217	-	-	-	-	F	-	-	-	-	-	-	-	-	F	-	-	-	L	M	F	-	-	-	-	-	-
consens.	-	R	-	I	-	-	-	Y	-	-	I	-	N	-	Y	-	D	L	-	-	D	-	-	S	N	-

Table 8.1. shows some of the key residues which contribute to the six pockets (A-F) within the peptide binding groove. The amino acid residues for each of the A*02 variants are compared to A*0201 at each of these positions, with substitutions denoted by the appropriate letter. A comparison is also made with a class I consensus sequence and A*0201 at the given amino acid positions. Of the alleles compared, only A*0209 is identical with A*0201 in all the residues considered.

important in binding to the HLA molecule, whilst others had influence on the interaction with the T cell receptor.

The influence of peptide sequence on binding to HLA class I was further explored by Madden (Madden, et al 1993). Five different viral peptides were co-precipitated with A*0201 and the conformation within the groove studied by X-ray crystallography. Each of the different peptides showed dramatic conformational changes, particularly in the centre of the peptide sequence. Hence in comparing the five different peptides, peptide positions P5, P6, P7 were shown to either point down into the groove, across towards the alpha helices or up towards the T cell receptor. This variation in peptide conformation was shown to inflict minor influence on the structure of A*0201. This suggests that the conformation of the peptide itself provides the dominant antigenic identity of the HLA class I complex in respect of the T cell receptor. It also suggests that peptide residues which are not involved in binding to the HLA molecule may be available to interact with the T cell receptor.

8.5.2. The influence of HLA polymorphism.

Most of the sixteen expressed allelic variants can be distinguished from each other by polymorphic differences found within the peptide binding groove (figure 8.2). Certain of these polymorphisms will have implications for the peptide anchor residues and so the binding specificity of the HLA-A*02 molecule. Others will exhibit influence on the conformation of peptide bound within the groove, and whether peptide residues point towards or away from the T cell receptor.

The effect of polymorphism on peptide binding and presentation has been investigated in a number of studies involving the use of mutagenesis (McMichael, et al 1988; Latron, et al 1991; Tussey, et al 1994). In one study, transplanting the B pocket of A*0201 into B*2705, caused the resulting class I molecule to present a different set of peptides compared to B*2705 with a resultant loss of B*2705 allorecognition (Colbert, et al 1993). This again demonstrated the importance of polymorphism within the peptide binding groove with regard to CTL recognition.

The positions of key amino acid residues within the class I molecule in the context of peptide binding and presentation has also been investigated (McMichael, et al 1988). CTL raised against the flu matrix peptide bound to A*0201 failed to lyse peptide pulsed A*02 variant donor cells containing A*0202, A*0203 and A*0205. Further to this, point mutations at amino acid positions 62, 63, 66, 152, 156 and in some instances 9 and 70 of the A*0201 molecule, caused loss of recognition by the A*0201 specific CTL clone. Mutations at positions 43, 74 and 107 had no effect on CTL recognition. In extending the study to examine another peptide in the context of A2 polymorphism, it became apparent that changes at positions 152, 155 and 156 did not abolish recognition in all cases. Hence these amino acid positions are more likely to determine the conformation of peptide within the class I molecule as determined by the peptide sequence (Tussey, et al 1994).

The relevance of the six peptide binding pockets in terms of CTL recognition of the flu matrix peptide was further investigated using a panel of mutants of HLA-A*0201 (Matsui, et al 1994). This study identified the importance of the B pocket in relation to presentation of this peptide. But it also showed the importance of other pockets in presentation of the peptide and CTL recognition. This study further underlines that even those pockets not involved in binding peptide anchor residues still have a critical role in orientation and presentation of peptide for CTL recognition (Matsui, et al 1994).

The determination of anchor residues, and so peptide motifs for a number of A*02 variants, has been established through several studies (Falk, et al 1991; Rotzschke, et al 1992; Barouch, et al 1995a; Sudo, et al 1995). In comparing the peptide motifs derived from A*0201 and A*0205, it became apparent that, although a number of epitopes presented to T cells would be similar, a considerable number would be different (Rotzschke, et al 1992). This was confirmed when the study was extended to additionally consider A*0202 and A*0214, alleles commonly found in parts of Africa (Barouch, et al 1995a). A further study examined peptides eluted from A*0204, A*0206 and A*0207, all of which contain a single amino acid change from A*0201 (Sudo, et al 1995).

Peptide motifs eluted in these particular studies further underlined the importance of amino acid position 9 in the HLA molecule. From Table 8.2, it can be seen that amongst the A*02 alleles, Phe_9 (A*0201, A*0202, A*0204, A*0207) is associated with a Leu P2 anchor, whilst Tyr_9 (A*0205, A*0206, A*0214) is associated with Val P2 anchor (Barouch, et al 1995a). This polymorphism at position 9 has relevance to the hydrophobic or hydrophilic nature of the B pocket, which effects the anchor residue bound at P2 (Barouch, et al 1995a). In A*0207, an anchor residue of Asp was found at P3 in addition to anchor residues of Leu at P2 and P9 (Sudo, et al 1995). The one amino acid substitution from A*0201 (Tyr_{99} to Cys_{99}), could potentially change the A*0207 D pocket in such a way that Asp at P3 could bind to a basic residue (His_{116}) in the bottom of the groove (Sudo, et al 1995).

The ability of different A*02 alleles to bind particular peptides has also been investigated. Assays which quantify binding of peptide to the class I molecule (Townsend, et al 1990; Tussey, et al 1994) or serologically defined conformational recognition of folded class I as a function of peptide binding (Fruci, et al 1993), have been used to identify which peptide motifs bind to which A*02 subtypes. Various studies have shown that peptides bound by one A*02 variant would cross bind to a number of other A*02 variants (Fruci, et al 1993; Tanigaki, et al 1994; del Guercio, et al 1995). It was therefore suggested that the A2 alleles formed a supertype which presented similar peptide motifs (del Guercio, et al 1995). This has been contradicted by others who have found that A*02 subtypes demonstrate distinct specificity for peptide epitopes (Barouch, et al 1995a; Sudo, et al 1995). This is further supported by the early observation that A*0202, A*0205 failed to present the A*0201 flu matrix peptide to influenza specific CTL (Gotch, et al 1985). Interestingly, in del Guercio's supertype hypothesis, A*0207 was excluded on the basis of its failure to bind certain A*02-type peptides (del Guercio, et al 1995). As discussed above, this may be explained by the peptide motif determined for A*0207 (Sudo, et al 1995). It seems that although several peptides can be bound by number of different A*02 variants, no one peptide is ubiquitous to all (Barouch, et al 1995a; Sudo, et al 1995). It has also been observed that some peptides seem to be specifically bound by only certain A*02 subtypes (Tussey, et al 1994). The level of crossreactivity of specific peptide bound appears to be a function of both peptide sequence and A*02 polymorphism.

	Position in peptide sequence									HLA polymorphism				
	P1	P2	P3	P4	P5	P6	P7	P8	P9	9	95	97	99	156
A*0201	-	L m	-	- ep	-	- ilv	-	-	L/V	F	V	R	Y	L
A*0202	-	L	-	-	-	-	-	-	L	-	L	-	-	W
A*0204	-	L	-	- ep	-	- il	-	-	L	-	-	M	-	-
A*0205	-	V ilq	-	-	-	- i	-	-	L	Y	L	-	-	W
A*0206	-	V q	- il	- ep	-	- l	-	-	V/L	Y	-	-	-	-
A*0207	-	L	D p	- e	-	- i	-	-	L	-	-	-	C	-
A*0214	-	V/Q l	-	-	-	- filv	-	-	L	Y	L	-	-	-

Table 8.2. A summary of peptide motifs determined for the given A*02 alleles, together with their key polymorphic differences. Letters in uppercase in the peptide sequence denote anchor residues. Letters in lowercase identify preferred residue at those positions within the peptide sequence.

A series of mutagenesis experiments has also demonstrated that polymorphism within A2 can cause presentation of endogenously processed peptide and not the equivalent exogenously added peptide. The reverse situation was also noted. This suggests that the mechanisms for class I assembly is different for presenting endogenously or exogenously derived peptides, and that these mechanisms are influenced by the A*02 mutant polymorphism (Matsui, et al 1995).

8.5.3. The role of CTL recognition.

The significance of A*02 polymorphism in terms of peptide presentation is further revealed in the specificity of HLA restricted CTL response. Certain CTL clones will recognise the same peptide bound to more than one A*02 variant. Other CTL clones differentiate between different A*02 variants binding the same peptide. Thus, both peptide and A*02 allele specificity contribute to T cell recognition, and different peptides in the context of A*02 polymorphism will exhibit different profiles of CTL clonal crossreactivity.
These different observations are supported by comparing binding of peptide with CTL recognition. This has been done in a comparison of A*0201, A*0202, A*0205 and A*0214 (Barouch, et al 1995a). The binding of HIV-1pol and flu matrix peptides were compared in the above four A*02 alleles. Both peptides bound well to A*0201 and A*0202, but less well to A*0205 and A*0214. When performing CTL recognition using a clone specific for each of the two peptides, the pol specific CTL recognised A*0201 and A*0202 but not A*0205 peptide pulsed cells, whilst the flu matrix specific CTL would only recognise this peptide bound to A*0201. The fact that a CTL can differentiate between two A*02 subtypes binding the same peptide suggests a change in peptide conformation as a result of A*02 polymorphism within the binding groove.

Hence, A*02 variants will exhibit certain differences in peptide binding through their preferred anchor residues (Table 8.2). Other sites of polymorphism within the peptide binding groove will determine the conformation of the peptide. This conformation will then determine which amino acid side chains within a particular peptide are accessible to the T cell receptor and so affect CTL recognition. Presentation of peptide to a T cell

receptor can therefore be seen as a function of peptide sequence and orientation as dictated by polymorphism within the HLA molecule to which it is bound.

8.6. Conclusion.

The dissection of HLA-A*02 as a group of closely related specificities has enabled a extended understanding of HLA polymorphism in a number of areas. The existance of HLA polymorphism can be conjectured as a response to selective pressures (Lawlor, et al 1990; Hill, et al 1991), and this seems borne out through the low level of silent nucleotide substitutions, the high level of polymorphism in the functionally sensitive peptide binding groove, and the conservation of polymorphic motifs in hypervariable regions through gene conversion events.

In analysis of HLA-A*02 polymorphism, the importance of even small changes in the peptide binding groove can be shown to have a marked effect on how the HLA molecule functions in presenting peptide to the T cell. As discussed, it seems likely that the generation of such polymorphism is in response to selective pathogenic pressure, as illustrated in the HLA-B53 restricted response to malaria in West Africa (Hill, et al 1991). It also seems probable that HLA diversity present in a particular population in part reflects the local history of pathogenic pressure.

Anthropological diversity can be illustrated amongst the variants of HLA-A*02, which are found at different frequencies in different populations (chapter 5, Table 5.2). Certain allelic variants seem prevalent in certain populations, which as mentioned may be in response to selective pressures. However the exact nature and diversity of A*02 alleles in different populations is still unclear due to the limited number of studies at this level of resolution. This lack of typing discrimination is certainly true for many class I alleles, but as typing resolution improves, so an increased understanding of HLA association with regard to anthropology, disease, transplantation and cellular function will follow.

The ability to type at allelic resolution has strengthened observations of linkage disequilibrium in different populations, as demonstrated with the A*0205 and B58 haplotype in Sardinia (Carcassi, et al 1995) and A*0207 and B46 in Singapore (Krausa, et al 1995). Disease association with HLA-A*02 has also become more apparent when considered at an allelic level of definition. Hence, an association has been made between A*0206 and Grave's disease, and between A*0207 and Hashimoto's thyroiditis in the Japanese population (Sudo, et al 1995).

HLA-A*02 provides an ideal allelic group to gauge the importance of polymorphism with respect to the functioning of the T cell mediated immune response. As additional studies address the role of HLA polymorphism, so a greater understanding of the mechanisms of immune interactions may provide positive solutions for the fields of transplantation, disease association and cellular immunology.

Chapter 9

Concluding Discussion

9.1 Aims & objectives for development of a PCR-SSP typing system at the HLA-A locus.

This thesis considers the definition of HLA-A locus alleles from DNA. The project was initiated at a time when HLA typing for class I specificities was securely within the domain of serology, with no practical DNA approach available. Serology had been firmly established as the method of choice for something approaching 30 years (Terasaki and McClelland, 1964). Over those years, serological definition of class I specificities had become increasingly refined as more was understood about HLA polymorphism through the improving interpretation of the specificity of antisera. Serology was nevertheless limited. It was not able to discriminate adequately between many class I specificities due to serological cross-reactivity which made the definition of certain specificities difficult, particularly when expressed in association with other class I specificities.

The application of molecular biological techniques to tissue-typing (Mullis, et al1986; Saiki, et al 1986; Saiki, et al 1989) initially provided no alternative to class I serology, since it was applied to the serologically impractical class II loci. HLA class II serology proved difficult through the requirement to isolate sufficient class II expressing B cells from the PBMC's. Also class II antisera had to be treated to remove the interference of class I specific antibodies. For these reasons amongst others, other methods of detecting

polymorphism were considered, including RFLP (Bell, et al 1987) and subsequently PCR-SSOP (Saiki, et al 1986), as discussed in chapter 1.

The analysis of class II polymorphism at the nucleotide level required the characterisation of all known specificities by DNA sequencing. This sequencing effort established expertise in defining polymorphism within DNA and led to an increased awareness of the complexity of polymorphism of HLA genes. It was inevitable that as more characterisation was directed at class II, so an increasing number of class I specificities would also become sequenced. However, serological tissue-typing for class I was considered adequate for most purposes, which diminished the incentive for development of DNA based approaches.

It has been apparent for some time, that serology was not sufficiently sensitive to identify all class I polymorphism, as observed in the HLA restricted discrimination of the T cell (Biddison, et al 1980; Gotch, et al 1985) or IEF analysis (Yang, 1989b) amongst other methods. However, these methods were not definitive in their own right and they provided no comprehensive alternative to serology. The relevance of serologically silent polymorphisms located around the peptide binding groove (Bjorkman, et al 1987) had implications as to the function of the class I molecule in binding and presenting antigen for immune surveillance. By not identifying these polymorphisms, serological typing was failing to discriminate between functionally distinct alleles.

In this thesis, the development of a DNA based HLA-A typing system provided an opportunity to identify these serologically silent yet functionally important polymorphisms within the HLA class I genes. The development of a DNA based method was permissible through the increased number of sequenced class I specificities. The specificity achievable through ARMS designed sequence specific primers (Newton, et al 1989) made feasible a simple PCR based approach, whereby the determination of polymorphism defining HLA specificity was resolved by the presence or absence of an amplification product. At the beginning of this thesis, this approach had been applied to definition of HLA class II specificities (Olerup and Zetterquist, 1991; Olerup and Zetterquist, 1992). The application of this approach with regard to HLA class I began simply, by identifying broad HLA-

A locus specificities. Through subsequent development, the system became increasingly refined through the use of more discriminatory primer combinations and the application of nested PCR, finally to provide a method which was capable of allele specific discrimination at the HLA-A locus (Krausa, et al 1996).

9.2. Comparison with other typing methods.

As discussed above, serology currently persists as the main approach to definition of class I specificities. At a practical level, DNA based methods hold many advantages over the serological approach to HLA typing. A significant limitation of serology lies in its reliance on viable PBMCs. Typings have to be done from fresh blood, or cells have to be suitably frozen to retain viability, if processing is required at a further date. The scarcity of good antisera, which are an increasingly limited and non-renewable resource, represents an additional restriction on the performance of serological typing. For this reason, different laboratories use different panels of antisera, which can contribute to typing inconsistencies through a lack of standardisation.

DNA based methods do not suffer from such restrictions. Good quality DNA can be extracted from a number of sources, including saliva, buccal scrapes and hair (Higuchi, et al 1988; Allen, et al 1994; Aron, et al 1994). DNA is easily stored and remains stable for long periods. This reduces the restrictions on sample transportation, increasing the scope of DNA approaches to type different population groups or archival material. DNA approaches can be standardised, particularly since oligonucleotides used as primers or probes, are synthetic and represent a renewable resource. Standardisation is also assisted by the use of conventional laboratory equipment in performing many of the DNA based techniques. A major advantage in applying DNA based approaches to HLA typing is that these methods offer a much higher resolution than that possible with serology.

The PCR-SSP method developed in this thesis requires a suitable PCR machine and gel electrophoresis equipment, items found routinely in many research and clinical laboratory settings. The widespread ability to perform

PCR facilitates the adoption of DNA based HLA typing methods into many laboratories.

A number of different DNA approaches have been developed in recent years and applied to detection of HLA polymorphism. There are four main generic methods of DNA typing currently used, although a range of variants for each have been described. These four general approaches (as reviewed in chapter 1), are (a) PCR-SSOP, (b) heteroduplex and SSCP analysis, (c) sequence based typing (SBT) and (d) PCR-SSP. Each method offers particular advantages which dictate their use for the appropriate applications. The attributes of the different approaches are discussed below.

(a) PCR-SSOP. PCR-SSOP has the advantage that large numbers of samples can be screened at each hybridisation event. However each sample requires hybridisation against a substantial number of probes which is time consuming. Situations may also occur using PCR-SSOP, when heterozygous combinations are difficult to interpret through the broadness of the initial PCR amplification. This can to some extent be rectified by performing a number of more specific PCR reactions (Fernandez-Viña, et al 1995; Middleton, et al 1995) so simplifying the interpretation of the hybridisation patterns.

PCR-SSOP can also detect and characterise the identification of new alleles. Novel hybridisation patterns may indicate new specificities. Because informative typing by PCR-SSOP requires screening with an extensive panel of oligonucleotide probes, the resulting characterisation of a new allele can provide a concise profile of its polymorphic content.

The use of PCR-SSOP through several past histocompatibility workshops has maintained and further developed this approach (Allen, et al 1994; Fernandez-Viña, et al 1995; Middleton, et al 1995; Gallardo, et al 1996). Such development has accumulated a wealth of experience in performing PCR-SSOP HLA typing, particularly in terms of the conditions crucial to obtaining consistent and accurate results. The communication of these experiences through workshops and publications, has allowed other laboratories to incorporate PCR-SSOP as one of their methods of histocompatibility testing. However, laboratories new to the field of molecular biology may find that

this inexperience represents a problem in performance of DNA based HLA typing techniques. For this reason, simplification and standardisation of the methods, making them accessible to a greater number of laboratories, has been an issue. The production of kits, particularly for the reverse dot-blot approach (Bugawan, et al 1994) (chapter 1), provides one approach for making a PCR-SSOP technique available for use in laboratories limited in molecular experience.

(b) Heteroduplex and SSCP analysis. Heteroduplex analysis and single stranded conformational polymorphism (PCR-SSCP) analyse the electrophoretic mobility of amplified DNA as determined by the nucleotide sequence. Heteroduplex analysis, as described in chapter 1, involves the denaturing and reanealing of PCR products, to create double stranded DNA combinations, not necessarily matched in terms of their sequence conformity. These novel combinations of DNA strands migrate differently in an electrophoresis gel as compared to a homoduplex, which can be applied to determining HLA sequence differences between a donor and recipient. However, this method does not determine the exact nature of a polymorphism in terms of its sequence. Hence it is a method of screening for polymorphism rather than for its specific definition in terms of HLA type.

Similarly, PCR-SSCP examines the migration of single stranded PCR product through a gel. The conformation of the single strand will be dependent upon its nucleotide sequence, which in turn determines its mobility through a gel under the appropriate conditions. This method again screens for polymorphism without identification of its exact nature in terms of its nucleotide sequence. Heteroduplex and SSCP analysis do not produce discriminatory results, and so fail to provide a comprehensive typing approach. They do however provide a useful method of screening for the presence of additional polymorphisms between seemingly matched individuals.

(c) Sequence based typing. In contrast to heteroduplex and PCR-SSCP analysis, the application of sequenced based typing (SBT) has the potential of providing the most definitive means of analysis of HLA polymorphism at the DNA level. SBT involves the direct sequencing of PCR products

amplified from regions of definitive polymorphism. By identifying all the given polymorphism within a region, this approach offers by far the most concise method of HLA typing. New alleles containing novel polymorphism within the regions sequenced will be definitively characterised by this method. The other molecular approaches may not always detect the presence of such polymorphism or may fail to provide a definitive characterisation as to their nature.

However, the SBT approach, although made easier through automated sequencing, is currently limited in terms of sample throughput and performance time. These limitations are being addressed through the application of new technology in terms of fluorescent labeled sequencing primers and sequencing equipment. Initial problems through the incorporation of Taq errors within the PCR prior to sequencing have also been dealt with by the use of improved proof reading enzymes. At present, SBT is not a replacement for PCR-SSOP and PCR-SSP, since these approaches have high capacity or speed respectively on their side. However, in situations were the most accurate method of typing is required, then SBT becomes highly attractive.

(d) PCR-SSP. PCR-SSP, which has been developed for typing at the HLA-A locus in this thesis, represents a simple and easy to interpret DNA based approach to tissue-typing. A considerable advantage of the PCR-SSP approach is the capability to perform a complete typing of a sample, at both class I and class II loci, under a standard set of conditions, in a relatively short period of time (Bunce, et al 1995b). This has obvious appeal in situations where a sample has to be characterised quickly. However, the number of samples that can be typed in any one period is limited in terms of the number of PCR amplifications that can be performed. The PCR-SSP approach is therefore in contrast to the sample batching and large throughput approach which characterises PCR-SSOP. PCR-SSP does hold similarities with the reverse dot-blot approach since both involve typing of a sample in a single event, be it against a panel of specific PCR reactions or hybridisation with a panel of oligonucleotide probes.

An advantage that PCR-SSP holds over PCR-SSOP is with regard to interpretation. With PCR-SSP, amplification of DNA requires the

determination of two points of polymorphisms along the length of its sequence as discriminated by the specificity of each primer. This in effect links two points of polymorphism present along a single stretch of sequence. Additionally, the use of nested PCR-SSP can characterise a given sequence on the basis of four points of linked polymorphism as determined by the specificity of each primer used in the initial and subsequent PCR. This is in contrast to PCR-SSOP, where each probing event is independent. The shared nature of polymorphism amongst HLA specificities, makes the interpretation of certain heterozygous combinations difficult by PCR-SSOP, since it becomes uncertain as to which polymorphisms are present together on which allelic sequence. Heterozygote combinations have not presented a problem for the PCR-SSP approaches developed in this thesis. As demonstrated in chapter 4.4.4, the ambiguity of defining certain A*02 heterozygote combinations was simply resolved through the addition of discriminatory PCR-SSP amplifications.

The requirement for the simultaneous performance of a large number of PCR amplifications has compelled PCR-SSP methods to embrace microtitre plate technology. This has facilitated easier performance of the typing technique, so allowing an increased throughput of samples. Microtitre plate formats are now available in 192 well and 384 wells per plate, which greatly increase the volume of typings and bring forward the possibility of automation (Merel, et al 1996). In addition, the basic philosophy of PCR-SSP can also be applied to testing for other polymorphic loci and these can be appended to the HLA typing panel. Hence, a single sample can be tested simultaneously for a whole array of polymorphism, in addition to the HLA loci.

New HLA sequences may be detected by PCR-SSP through a novel pattern of reactivity with the reaction panel. Conversely, PCR-SSP is limited in as much as it may fail to detect novel sequence polymorphisms which cannot be determined within the discrimination of the existing panel of primer combinations. Hence, a novel HLA sequence, which is functionally discreet, may be misassigned. PCR-SSP can only be completely reliable within the context of the known sequences. This is also true for PCR-SSOP. Typing by more than one typing method allows a sample to be gauged from a number of different perspectives. Any inconsistency between different typing

methods may herald the discovery of a new HLA specificity. SBT has a distinct advantage in the detection of new HLA specificities, since all the polymorphism is sequenced and determined within a given region of interest.

As described in chapter 6, several new HLA-A specificities have been determined through the course of this thesis. Three of the four alleles described presented as anomalies in the expected typing result as a consequence of their novel sequence recombination. The fourth allele was characterised because of a discrepancy between methods, and would not have been determined with the existing PCR-SSP panel, another endorsement of a SBT approach. The use of PCR 'gene-mapping' was applied to the three anomalous typing results. This method provided rapid verification of the observed reaction patterns seen in the new alleles, and correctly predicted their polymorphic content (as confirmed by conventional DNA sequencing) within the context of SSPs defining known sequence motifs. In this thesis, PCR-SSP is shown to detect and characterise new alleles, but situations may occur when a novel specificity is either misassigned or not detected.

The ability to tissue-type from DNA is now well established, and is being swept along with the innovations increasingly offered in the field of molecular biology. The PCR-SSP approach developed in this work clearly demonstrated the feasibility of tissue-typing at the HLA-A locus at a time when molecular approaches were not widely regarded for class I specificities. Additionally, it showed the ability for high resolution allelic definition, so demonstrating the advantages of DNA based approaches over serological testing. On occasion, typing based on surface expression of HLA molecules does have an advantage, particularly if DNA based approaches fail to discriminate alleles which are not expressed, detected as blanks by serology. Several such instances, as described in chapter 6 do exist but appear to be rare. However once the sequence polymorphism which identifies them is determined, then DNA based methods again can be used for their identification.

The developments from this thesis provided an alternative approach to serology at the HLA-A locus. Work at other loci determined that a complete

PCR-SSP approach was feasible for both class I and class II detection (Olerup and Zetterquist, 1992; Olerup, et al 1993; Bunce, et al 1994; Sadler, et al 1994; Bunce, et al 1995a). The developments made at the HLA-A locus in this thesis together with work at other loci, has contributed to a PCR-SSP method for typing HLA class I in the 12th International Histocompatibility Workshop and for class I and class II in work by Bunce et al (Bunce, et al 1995b).

The application of different techniques to the field of HLA testing has generated a battery of methods, each with their perceived advantages and disadvantages, each with their supporters and critics. The reasons why a sample is HLA typed will determine to some extent the method by which this is achieved. Factors such as speed, volume, technical expertise, cost, required resolution amongst others, further identify the most appropriate approach. Serology has been the method of choice for class I typing for many years. It is unlikely in view of the strides made in development of DNA approaches whether the dominance of serology will persevere (Dupont, 1995). Given the relative merits of different methods in different applications, no single approach may emerge as the method of choice, although the high level of discrimination offered by SBT suggests it may become the "gold standard" for high definition typing. As new DNA based methods emerge, and knowledge increases as to the significance of HLA polymorphism in the immune function, it is unclear as to which approaches will be adopted for the different applications of histocompatibilty testing.

9.3. The applications of the PCR-SSP approach.

HLA typing represents an important aspect for a number of fields, including transplantation, population studies, cellular immunology and disease association. Following the development of the PCR-SSP system for HLA-A locus typing, a number of studies presented themselves in which a high resolution typing approach would be both informative and useful. Some of these studies have been described in the preceding chapters. In detail, these include population studies (chapter 5), characterisation of new HLA specificities (chapter 6), significance in cellular immunology (chapters 7 and

8). In addition, the method has been used to investigate disease association and this will be commented on below.

(a). Characterising new alleles. The application of PCR-SSP to the typing of HLA-A locus specificities created a novel perspective from which a number of new alleles were identified. The issues involved in the identification of new HLA-A alleles are discussed in full in chapter 6. The development of a PCR-SSP approach for HLA-A locus provided a realisation as to the sensitivity of the technique in defining HLA polymorphism. Through the specificity of ARMS (Newton, et al 1989) designed SSP, it was possible to differentiate between two sequences on the basis of a single nucleotide difference. This power of this discrimination provided the basis of the developed PCR-SSP system.

The possible discrimination also offered a means of rapidly analysing sequences within the context of known SSP sites. This was facilitated through the shared nature of polymorphism and relative lack of unique sequence substitutions, found amongst class I alleles. This was the basis of the development of the PCR gene-mapping technique (chapter 6.2), whereby a panel of SSP could define the sequence motifs present within a particular allele, in the context of known polymorphism. This method successfully predicted the sequence of a number of new alleles, as described in chapter 6 and elsewhere (Browning, et al 1995). The method fails to predict a sequence where the polymorphism is not defined by the SSP panel. Characterisation of new alleles by this method proved rapid, and allowed the quick design of new SSP combinations for the determination of the new specificity. It also allowed the evaluation of the typing panel specificity with regard to the new allelic sequence. Samples which potentially contained the new specificity could be quickly screened with the appropriate SSP combinations. New alleles defined by PCR-SSP still required full length sequencing, to verify the polymorphism present, and confirm the sequence in regions not characterised by the gene-mapping approach.

The application of PCR gene-mapping, proved reliable for the sequence characterisation of alleles tested in chapter 6. But in performing gene-mapping, it is important to be aware of the limitations of the method, particularly with regard to the analysis of alleles containing novel

polymorphisms (e.g. A*8001) which may fall outside of the hypervariable regions of the class I sequence (e.g. A*0209, A*0215N). SBT offers a more substantive approach to detection and characterisation of a new allele. But even this technique may be inadequate in characterising a new allele, if the distinguishing polymorphism falls outside of the sequencing template.

As discussed in chapter 1, the number of HLA alleles has increased dramatically with the expansive application of molecular biological techniques. In examining the polymorphism amongst this increasing database of HLA sequences, it is clear that a mechanism such as gene conversion (Parham, et al 1988) has contributed to a large proportion of the diversity seen within the HLA system. To this end, the extent of HLA diversity has probably not been fully determined, and there is considerable scope for identification of new alleles generated from new gene conversion events. PCR-SSP may identify such alleles through observations made in typing, and the approach of PCR gene-mapping provides a rapid means for their preliminary characterisation. The application of DNA based typing techniques to different population groups may also uncover new alleles, as demonstrated in this thesis with respect to A*3004 (Sardinia) and A*0214 (Kenya) (chapter 6.3).

(b). Population studies. Examples of serological characterisation of HLA class I specificities in different populations have been reported in a number of studies (duToit, et al 1990; Allsopp, et al 1992; Contu, et al 1992; Halle, et al 1994), and as part of Histocompatibility Workshops (Imanishi, et al 1992a; Imanishi, et al 1992b). Serology does reveal differences between population groups, but fails to show this to its full extent, through a lack of resolution. The application of a high resolution typing approach as detailed in this thesis provided the opportunity for a more precise analysis of HLA within a population.

In determining HLA polymorphism within a population, particularly through family studies it is evident that certain HLA specificities are inherited together as part of a haplotype. This was demonstrated in a Sardinian study (Contu, et al 1992), were A*30 was found to be within one of the major haplotypes within the population. Nested PCR-SSP subtyping was performed on A*30 individuals from this population (chapter 5.2) to

determine if any particular A*30 allele constituted a member of this haplotype. The results indicated that A*3002 was the allele in the haplotype, greatly strengthening this association. In other studies (chapter 5), A*0205 was found to be the allele in an A2-B17 haplotype observed in the Sardinian population. Similarly A*0207 was found to be in linkage disequilibrium with B*46, strengthening an A2-B46 association noted in the Singapore Chinese population. These associations had initially been made through serological observations and were weakened through lack of resolution. The application of a high resolution DNA approach was shown to focus the association between HLA specificities, so strengthening the relationship.

The application of high resolution PCR-SSP provided a means for comparison of HLA-A*02 subtypes in a number of different population groups (chapter 5.3). A*02 is commonly found in many different populations, and this study showed that the A*02 allelic composition varied greatly between the groups investigated. Hence, the resolution offered by PCR-SSP determined heterogeneity between populations, which was not apparent through serological analysis.

PCR-SSP was also applied to typing a Ugandan population (chapter 5.4), using both the PCR-SSP Phototyping method (Bunce, et al 1995b), with additional high resolution typing performed by nested PCR. The serological determination of many of the specificities identified, would have proven difficult, through a lack of antisera determining specificities both peculiar to and common in African populations. Such specificities were readily identified by PCR-SSP. On a practical note, serological analysis of this population would have proven restrictive through the requirement for viable cells, a constraint not made on the DNA approach. The application of high resolution HLA typing revealed a more defined HLA profile than that possible with serology. The population differences demonstrated by high resolution typing have major implications for the fields of transplantation and cellular immunology with regard to matching between groups. In addition, the high resolution analysis of HLA in different populations may provide an improved anthropological insight into the relationship between different ethnic groups.

(c). Disease association. This application of the PCR-SSP HLA-A typing approach to disease association was not addressed as a study in this thesis. However, through the development of the method, various studies have benefited from the typing system through a number of collaborations. One such collaboration involved the typing part of a cohort of commercial sex workers, attending a clinic in Nairobi (Plourde, et al 1992). It was noted that some of the women attending the clinic failed to become HIV infected, despite a history of multiple contact with HIV positive partners. An initial serological definition of this population proved inaccurate. The application of PCR-SSP typing at A locus provided more definitive results. Unfortunately no clear correlation was observed at the HLA-A locus with non-infection with HIV.

Another collaborative study examined the relationship between HLA and scarring trachoma in The Gambia in West Africa. Infected individuals were serologically HLA typed and an association was noted between HLA-A28 and contracting scarring trachoma as compared to the control group. The application of PCR-SSP both removed mistypings (A28 is difficult to determine by serology, particularly in the presence of A2), and showed that amongst the A28 samples, the association was clearly with the A*6802 subtype (Conway, 1996).

Both the above studies exemplified the inadequacies of serology in providing accurate data for determination of a disease association. The application of high resolution DNA based approaches as illustrated, provide more informative data from which the presence of any associations could be determined.

(d). HLA and functional significance. In the latter part of this thesis, the importance of allelic polymorphism was both explored and discussed (chapters 7 and 8). The purpose of this thesis was to determine HLA-A locus specificities to an allelic level. These two chapters contemplate the importance of that polymorphism with regard to the functioning of the HLA molecule in the MHC restricted immune response.

Chapter 7 examines the relevance of HLA polymorphism in respect to the binding of peptide for presentation to the TCR. Unfortunately, it was not

possible to determine the contribution made by the TCR in this study, although this will be addressed in future work. However, the data obtained from eluting peptides from variants of A*30, showed a difference in anchor residues preference in A*3001 as opposed to the other variants. The significance of other A*30 variant polymorphic differences not involved in binding peptide anchor residues was also discussed in chapter 7, although the conclusions drawn can only be validated through the discrimination of the TCR.

Chapter 8 reviewed polymorphism in HLA-A*02 and its relationship to function. As described, a number of studies have examined and showed the relevance of polymorphism amongst various A*02 alleles with regard to the T cell response. The observations made in chapter 7 and 8 and their implications provide validation of HLA typing to a high resolution, as a means of accurately determining functionally distinct molecules at the cellular level. The importance of HLA typing to the allelic level is therefore a prerequisite for the confident interpretation of MHC based cellular interactions.

9.4. The future.

The demands for tissue typing remain through the requirements of transplantation, forensics, studies in disease association, anthropology and T-cell mediated immunity. Each of these areas has different requirements as to the degree of resolution and methodology, which explains the variety of techniques available for HLA tissue typing. The advent and rapid progress of DNA methodology has provided a workable alternative to serology, although serology will probably remain the routine method for HLA class I typing in many HLA laboratories for the near future.

The practice of serology will inevitably decline as the opportunity for DNA typing reaches an increasing number of laboratories, through the vehicles of workshops, and commercial kits. The availability of DNA based kits, which require only standard laboratory equipment, will potentially take tissue typing out of the realm of the professional histocompatibility testing laboratory, into the hands of anyone who has a reason for doing it. Increased automation may further reduce specialised knowledge. However, the

complexity of the HLA system and diversity of available methods should ensure that the requirement for professional histocompatibility testing laboratories persists for the near future.

The information generated by DNA based methods has contributed to a more precise understanding of the HLA system from a number of different perspectives. For instance, HLA typing at a high resolution has uncovered previously unseen heterogeneity between population groups in terms of their comparative structure of HLA gene frequencies (see chapter 5). DNA based HLA typing will hopefully improve the area of transplantation by identifying more precisely which specificities or combinations of sequence polymorphisms are crucial for matching to ensure graft survival. With many of the HLA polymorphisms having functional significance at the level of the T cell mediated immune response, high resolution typing also has implications for interpretation of HLA restriction in both research and clinical settings. For these reasons, the requirement for precise and dependable typing methods will continue to ensure further progress and development in the field of HLA typing.

The current advances and improvement of typing resolution, has generated an extensive effort in further characterising polymorphism within the HLA system. As demonstrated in this thesis, new alleles are constantly being found and their functional and anthropological significance can subsequently be investigated. The information gained from such efforts will enhance current understanding of the structure, function and evolution of the HLA system. The diversity of methods currently available for HLA typing, provides the investigator a degree of choice, flexibility and accuracy not previously known in HLA typing.

9.5 In summary.

This thesis demonstrates the feasibility and subsequent development of a molecular typing of HLA class I polymorphism. The PCR-SSP method developed is a straightforward and easy to interpret approach to the definition of HLA polymorphism. The method was developed to allow a considerable number of sequence specific reactions to be performed

simultaneously under a set of standard conditions. The presence of a sequence is defined simply by its amplification by one or more of a panel of PCR reactions. By visualising the panel of PCR amplifications, a tissue type can be determined simply by the presence or absence of product in each of the panel of specific reactions. This thesis describes the development of a PCR-SSP approach as applied to defining specificities at the HLA-A locus.

Establishing that molecular approaches could be applied at the class I HLA-A locus demonstrated that an alternative existed to serological definition. In addition the advantages of typing from DNA were realised since viable cells were no longer necessary. Distant population groups, as described in chapter 5, could be typed without the inherent restrictions accompanying serological definition. Combinations of specificities which serological typing finds difficult were easily and distinctly defined by PCR-SSP.

The major advantage of typing from DNA lies in the high resolution possible. Serology is limited in that it applies only to the definition of those polymorphisms which are either accessible or effect sites that are accessible to detection by antibody. Analysis of the nucleotide sequence provides the opportunity to analyse all the sites of polymorphism present within a HLA specificity, plus intronic sequences and promoter polymorphism. The ability to differentiate between specificities upon the basis of their sequence proffers the possibility of discriminating to a level of resolution at which the allele can be unambiguously identified. This thesis has demonstrated the application of PCR-SSP techniques, in defining HLA-A locus specificities well beyond the resolution of serology, so as to be able to type definitively to the level of the allele.

In a clinical or cellular setting, it may not be necessary to discriminate between all alleles, as some may contain polymorphisms that have no functional impact. However, for this to be established requires initial analysis based on allelic definition. In situations in which HLA is used as a marker, for instance in anthropology or disease association studies, then definition at the allelic level is more appropriate. As molecular HLA typing techniques become increasingly robust, easier to perform and standardised, then it will be convenient to type at the highest level of resolution for all

applications since this will provide a universal and standard form of information.

The necessity of typing to the point of defining the allele is addressed in chapters 7 and 8 of this thesis. The definition of HLA polymorphism and its study in comparison to function, provides a valuable insight into the mechanisms shaping the MHC restricted immune response. As investigated in this thesis, the appropriate HLA polymorphism determines the preference for residues bound within a peptide sequence. In addition, polymorphism will also effect how peptide lies within the binding groove and which residues provide a surface in the context of the class I molecule for interaction with the T cell receptor. The functional significance of polymorphism which distinguish the majority of class I alleles provide strong reasons for their accurate identification and so the use of a definitive typing method.

This thesis has therefore addressed the issue of providing an accurate and powerful DNA based system for definition and discrimination of alleles of the HLA-A locus. It describes strategies which simply and clearly facilitate allelic resolution. The approaches developed in this work have been adopted and validated in other laboratories (Bozón, et al 1996), and found to be an accurate and practical approach to class I DNA based typing. This thesis therefore marks an important point in the field of histocompatibility testing, demonstrating the feasibility and importance of PCR-SSP, as a means of providing high resolution typing for alleles of the HLA-A locus.

References.

Aidoo, M., Lalvani, A., Allsopp, C.E., Plebanski, M., Meisner, S.J., Krausa, P., Browning, M., Morris, J.S., Gotch, F., Fidock, D.A. and et, a.l. Identification of conserved antigenic components for a cytotoxic T lymphocyte-inducing vaccine against malaria. *Lancet 345:* 1003-7, 1995.

Allen, M., Liu, L. and Gyllensten, U. A comprehensive polymerase chain reaction-oligonucleotide typing system for the HLA class I A locus. *Hum Immunol 40:* 25-32, 1994.

Allsopp, C.E., Harding, R.M., Taylor, C., Bunce, M., Kwiatkowski, D., Anstey, N., Brewster, D., McMichael, A.J., Greenwood, B.M. and Hill, A.V. Interethnic genetic differentiation in Africa: HLA class I antigens in The Gambia. *Am J Hum Genet 50:* 411-21, 1992.

Allsopp, C.E., Hill, A.V., Kwiatkowski, D., Hughes, A., Bunce, M., Taylor, C.J., Pazmany, L., Brewster, D., McMichael, A.J. and Greenwood, B.M. Sequence analysis of HLA-Bw53, a common West African allele, suggests an origin by gene conversion of HLA-B35. *Hum Immunol 30:* 105-9, 1991.

Amos, D.B. and Yunis, E.J. Editorial: A new interpretation of the major histocompatibility gene complexes of man and mouse. *Cell Immunol 2:* 517-20, 1971.

Ariyoshi, K., Cheingsong, P.R., Wilkins, A., Corrah, T., Weber, J. and Whittle, H. HIV-1 subtype B in west Africa [letter]. *Lancet 347:* 1996.

Arnett, K.L. and Parham, P. HLA Class I nucleotide sequences, 1995. *Tissue Antigens 46:* 217-257, 1995.

Bach, F.H. and Amos, D.B. Hu-1: Major histocompatibility locus in man. *Science 156:* 1506-8, 1967.

Bach, F.H. and Voynow, N.K. One-way stimulation in mixed leukocyte cultures. *Science 153:* 545-7, 1966.

Balas, A., S. F. Garcia, R. F. Gomez and J. L. Vicario. HLA class I allele (HLA-A2) expression defect associated with a mutation in its enhancer B inverted CAT box in two families. *Hum Immunol* 41(1): 69-73: 1994.

Barber, D.F., Fernandez, J.M., Guttridge, M.G. and Lopez, d.C.J. Primary structure of a new HLA-A2 subtype: HLA-A*0213. *Immunogenetics 39:* 1994.

Barber, L., T. Patel, L. Percival, J. Gumperz, L. Lanier, J. Phillips, J. Bigge, M. Wormwald, R. Parekh and P. Parham. Unusual uniformity of the N-linked oligosaccharides of HLA-A, -B, and -C glycoproteins. *J Immunol* 156(9): 3275-84: 1996.

Barouch, D., Friede, T., Stevanovic, S., Tussey, L., Rowland-Jones, S., Braud, V., McMichael, A.J. and Rammensee, H.G. HLA-A2 subtypes are functionally distinct in peptide binding and presentation. *Journal of Experimental Medicine 182:* 1847-1856, 1995a.

Barouch, D., Krausa, P., Bodmer, J., Browning, M.J. and McMichael, A.J. Identification of a novel HLA-A2 subtype, HLA-A*0216. *Immunogenetics 41:* 388-388, 1995b.

Barnstable, C., W. Bodmer, G. Brown, G. Galfre, C. Milstein, A. Williams and A. Ziegler. Production of monoclonal antibodies to group A erythrocytes, HLA and other human cell surface antigens-new tools for genetic analysis. *Cell* 14(1): 9-20: 1978.

Bein, G., Glaser, R. and Kirchner, H. Rapid HLA-DRB1 genotyping by nested PCR amplification. *Tissue Antigens 39:* 68-73, 1992.

Bell, J. The polymerase chain reaction. *Immunol Today 10:* 351-5, 1989.

Bell, J.I., Denney, D.J., MacMurray, A., Foster, L., Watling, D. and McDevitt, H.O. Molecular mapping of class II polymorphisms in the human major histocompatibility complex. I. DR beta. *J Immunol 139:* 562-73, 1987.

Biddison, W.E., Kostyu, D.D., Strominger, J.L. and Krangel, M.S. Delineation of immunologically and biochemically distinct HLA-A2 antigens. *J Immunol 129:* 730-4, 1982.

Biddison, W.E., Krangel, M.S., Strominger, J.L., Ward, F.E., Shearer, G.M. and Shaw, S. Virus-immune cytotoxic T cells recognize structural differences between serologically indistinguishable HLA-A2 molecules. *Hum Immunol 1:* 225-32, 1980.

Bidwell, J.L., Bidwell, E.A., Laundy, G.J., Klouda, P.T. and Bradley, B.A. Allogenotypes defined by short DQ alpha and DQ beta cDNA probes correlate with and define splits of HLA-DQ serological specificities. *Mol Immunol 24:* 513-22, 1987.

Bidwell, J.L., Bidwell, E.A., Savage, D.A., Middleton, D., Klouda, P.T. and Bradley, B.A. A DNA-RFLP typing system that positively identifies serologically well-defined and ill-defined HLA-DR and DQ alleles, including DRw10. *Transplantation 45:* 640-6, 1988.

Birnboim, H.C. and Doly, J. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Research 7:* 1513-1522, 1979.

Bjorkman, P.J., Saper, M.A., Samraoui, B., Bennett, W.S., Strominger, J.L. and Wiley, D.C. Structure of the human Class I histocompatibility antigen, HLA-A2. *Nature 329:* 506-512, 1987.

Bodmer, J. Ia antigens: definition of the HLA-DRw specificities. *Br Med Bull* 34(3): 233-40: 1978.

Bodmer, J., Bodmer, W., Heyes, J., So, A., Tonks, S., Trowsdale, J. and Young, J. Identification of HLA-DP polymorphism with DP alpha and DP beta probes and monoclonal antibodies: correlation with primed lymphocyte typing. *Proc Natl Acad Sci U S A 84:* 4596-600, 1987.

Bodmer, J., S. Marsh and J. Heyes. Monoclonal antibodies and the epitope distribution of HLA class II polymorphisms. *Cold Spring Harb Symp Quant Biol* 54 Pt 1: 471-6: 1989.

Bodmer, J.G., Marsh, S.G., Parham, P., Erlich, H.A., Albert, E., Bodmer, W.F., Dupont, B., Mach, B., Mayr, W.R., Sasazuki, T. and et, a.l. Nomenclature for factors of the HLA system, 1989. *Tissue Antigens 35:* 1-8, 1990.

Bodmer, J.G., Marsh, S.G.E., Albert, E.D., Bodmer, W.F., Dupont, B., Erlich, H.A., Mach, B., Mayr, W.R., Parham, P., Sasazuki, T., Schrueder, G.M.T., Strominger, J.L., Svejgaard, A. and Terasaki, P.I. Nomenclature for factors of the HLA system, 1994. *Tissue Antigens 44:* 1-18, 1994.

Bodmer, J.G., Marsh, S.G.E., Albert, E.D., Bodmer, W.F., Bontrop, R.E., Charron, D.C., Dupont, B., Erlich, H.A., Mach, B., Mayr, W.R., Parham, P., Sasazuki, T., Schrueder, G.M.T., Strominger, J.L., Svejgaard, A. and Terasaki, P.I. Nomenclature for factors of the HLA system, 1995. *Tissue Antigens 46:* 1-18, 1995.

Bodmer, J., P. Pickbourne, et al. Serological identification of Ia antigens: report of a British region Ia workshop. *Tissue Antigens* 8(5): 359-71: 1976.

Bodmer, W.F., Albert, E., Bodmer, J.G., Dausset, J., Kissmeyer, N.F., Mayr, W., Payne, R., van, R.J., Trnka, Z. and Walford, R.L. Nomenclature for factors of the HLA system 1984. *Immunogenetics 20:* 593-601, 1984.

Bodmer, W.F. and Bodmer, J.G. (1979) Cytofluorochromasia for HLA-A, -B, -C and DR typing. In: NIAID manual of tissue typing techniques 1979-1980. NIH. (eds), p46-54 Bethesda, Maryland.

Bodmer, W. F., M. Tripp and J. Bodmer (1967). Application of a fluoro-chromatic cytotoxicity assay to human leukocyte typing. in; Histocompatibility Testing 1967. E. S. Curtoni, P. L. Mattiuz and R. M. Tosi. Copenhagen, Munksgaard: 341-350.

Bozón, M. V., Delgado, J. C., Turbay, D., Salazar, M., Granja, C. B., Alosco, S. M., Dupont, B. and Yunis, E. J. Comparison of HLA-A antigen typing by serology with two polymerase chain reaction based DNA methods: implications for proficiency testing. *Tissue Antigens* 47: 512-518, 1996.

Brodsky, F.M., Lem, L. and Bresnahan, P.A. Antigen processing and presentation. *Tissue antigens 47:* 464-471, 1996.

Brodsky, F., P. Parham, C. Barnstable, M. Crumpton and W. Bodmer. Monoclonal antibodies for analysis of the HLA system. *Immunol Rev* 47: 3-61: 1979.

Brown, J.H., Jardetzky, T.S., Gorga, J.C., Stern, L.J., Urban, R.G., Strominger, J.L. and Wiley, D.C. Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1 [see comments]. *Nature 364:* 33-9, 1993.

Browning, M.J., Krausa, P. and Bodmer, J.G. Low resolution HLA-A locus typing by multiplex ARMS-PCR [abstract]. *Human Immunology 39:* 134, 1994.

Browning, M.J., Madrigal, J.A., Krausa, P., Kowalski, H., Allsopp, C.E.M., Little, A.M., Turner, S., Adams, E.J., Arnett, K.L., Bodmer, W.F., Bodmer, J.G. and Parham, P. The HLA-A, -B, -C genotype of the class I negative cell line Daudi reveals novel HLA-A and -B alleles. *Tissue Antigens 45:* 177-187, 1995.

Browning, M., Petronzelli, F., Bicknell, D., Krausa, P., Rowan, A., Tonks, S., Murray, N., Bodmer, J. and Bodmer, W. Mechanisms for loss of HLA class I expression on colorectal tumor cells. *Tissue Antigens 47:* 364-371, 1996.

Buffone, G.J. and Darlington, G.J. Isolation of DNA from biological specimens without extraction with phenol [letter]. *Clin Chem 31:* 164-5, 1985.

Bugawan, T.L., Apple, R. and Erlich, H.A. A method for typing polymorphism at the HLA-A locus using PCR amplification and immobilized oligonucleotide probes. *Tissue Antigens 44:* 137-47, 1994.

Bunce, M., Barnardo, M.C. and Welsh, K.I. Improvements in HLA-C typing using sequence-specific primers (PCR-SSP) including definition of HLA-Cw9 and Cw10 and a new allele HLA-"Cw7/8v". *Tissue Antigens 44:* 200-3, 1994a.

Bunce, M., Fanning, G.C. and Welsh, K.I. Comprehensive, serologically equivalent DNA typing for HLA-B by PCR using sequence-specific primers (PCR-SSP). *Tissue Antigens 45:* 81-90, 1995a.

Bunce, M., O'Neill, C.M., Barnardo, M.C.N.M., Krausa, P., Browning, M.J., Morris, P.J. and Welsh, K.I. Phototyping: Comprehensive DNA typing for HLA-A, B, C, DRB1, DRB3, DRB4, DRB5 & DQB1 by PCR with 144 primer mixes utilising sequence-specific primers (PCR-SSP). *Tissue Antigens 46:* 355-367, 1995b.

Bunce, M., Taylor, C.J. and Welsh, K.I. Rapid HLA-DQB typing by eight polymerase chain reaction amplifications with sequence-specific primers (PCR-SSP). *Hum Immunol 37:* 201-6, 1993.

Bunce, M. and Welsh, K.I. Rapid DNA typing for HLA-C using sequence-specific primers (PCR-SSP): identification of serological and non-serologically defined HLA-C alleles including several new alleles. *Tissue Antigens 43:* 7-17, 1994.

Carcassi, C., Krausa, P., Bodmer, J., Contu, L. and Browning, M. Characterization of HLA-A*02 subtypes in the Sardinian population. *Tissue Antigens 46:* 391-393, 1995.

Carpenter, C. B., P. Dyer, et al. (1989). Antigen Society #3 Report (AW19 and A32/Aw74). in; Immunobiology of HLA. Histocompatibility testing 1987. B. Dupont. New York, Springer-Verlag. 1: 124-126.

Castano, A.R., Lauzurica, P., Domenech, N. and Lopez, d.C.J. Structural identity between HLA-A2 antigens differentially recognized by alloreactive cytotoxic T lymphocytes. *J Immunol 146:* 2915-20, 1991.

Ceppellini, R. and Rood, J.v. The HL-A system. I. Genetics and molecular biology. *Semin Hematol 11:* 233-51, 1974.

Cereb, N., Lee, S., Maye, P., Kong, Y. and Yang, S.Y. Nonrandom allelic variation in the regulatory complex of HLA class I genes. *Hum Immunol 41:* 46-51, 1994.

Cereb, N., Maye, P., Lee, S., Kong, Y. and Yang, S.Y. Locus-specific amplification of HLA class I genes from genomic DNA: locus-specific sequences in the first and third introns of HLA-A, -B, and -C alleles. *Tissue Antigens 45:* 1-11, 1995.

Chen, D.F., Endres, W., Meyer, S.A. and Stangel, W. A polymerase chain reaction-sequence-specific oligonucleotide procedure for HLA class II typing using biotin- and digoxigenin-labeled probes simultaneously in hybridization. *Hum Immunol 39:* 25-30, 1994.

Choo, S.Y., Starling, G.C., Anasetti, C. and Hansen, J.A. Selection of an unrelated donor for marrow transplantation facilitated by the molecular characterization of a novel HLA-A allele. *Hum Immunol 36:* 20-6, 1993.

Clay, T.M., Bidwell, J.L., Howard, M.R. and Bradley, B.A. PCR-fingerprinting for selection of HLA matched unrelated marrow donors. Collaborating Centres in the IMUST Study. *Lancet 337:* 1049-52, 1991.

Clay, T.M., Culpan, D., Howell, W.M., Sage, D.A., Bradley, B.A. and Bidwell, J.L. UHG crossmatching. A comparison with PCR-SSO typing in the selection of HLA-DPB1-compatible bone marrow donors. *Transplantation 58:* 200-7, 1994.

Colbert, R.A., Rowland, J.S., McMichael, A.J. and Frelinger, J.A. Allele-specific B pocket transplant in class I major histocompatibility complex protein changes requirement for anchor residue at P2 of peptide. *Proc Natl Acad Sci U S A 90:* 6879-83, 1993.

Contu, L., Arras, M., Carcassi, C., La Nasa, G. and Mulgaria, M. HLA structure of the Sardinian population: A haplotype study of 551 families. *Tissue Antigens 40:* 165-174, 1992.

Conway, D.J., Holland, M.J., Cambell, A.E., Bailey, R.L., Krausa, P., Peeling, R., Whittle, H.C. and Mabey, D.C.W. HLA class I and II polymorphisms and trachomatous scarring in Chlamydia trachomatis - endemic population. *Journal of Infectious Diseases 174:* 643-646., 1996.

Date, Y., Kimura, A., Kato, H. and Sasazuki, T. DNA typing of the HLA-A gene: population study and identification of four new alleles in Japanese. *Tissue Antigens 47:* 93-101, 1996.

Dausset, J. Leuco-agglutinins. IV. Leuco-agglutinins and blood transfusion. *Vox Sanguis 4:* 190-198, 1954.

Davis, M. and P. Bjorkman. T-cell antigen receptor genes and T-cell recognition [published erratum appears in Nature 1988 Oct 20;335(6192):744]. *Nature* 334(6181): 395-402: 1988.

Degli-Esposti, M.A., Griffiths, M.S., Daly, L.N., Witt, C.S., Simons, M., Carcassi, C., Albert, E.D., Giphart, M.J. and Dawkins, R.L. Characterization of 4AOHW cell line panel including new data for the 10IHW panel. *Hum Immunol 38:* 3-16, 1993.

del Guercio, M., Sidney, J., Hermanson, G., Perez, C., Grey, H.M., Kubo, R.T. and Sette, A. Binding of a peptide antigen to multiple HLA alleles allows definition of an A2-like supertype. *J Immunol 154:* 685-93, 1995.

de Preval, C. and Mach, B. The absence of beta 2-microglobulin in Daudi cells: active gene but inactive messenger RNA. *Immunogenetics 17:* 133-40, 1983.

Dieye, A., Diaw, M.L., Rogier, C., Trape, J.F. and Sarthou, J.L. HLA-A, -B, -C, -DR, -DQ typing in a population group of Senegal: distribution of HLA antigens and HLA-DRB1*13 and DRB1*11 subtyping by PCR using sequence specific primers (PCR-SSP). *Tissue Antigens 47:* 194-199, 1996.

Domena, J.D., Hildebrand, W.H., Bias, W.B. and Parham, P. A sixth family of HLA-A alleles defined by HLA-A*8001. *Tissue Antigens 42:* 156-159, 1993.

Domena, J.D., Little, A.M., Arnett, K.L., Adams, E.J., Marsh, S.G. and Parham, P. A small test of a sequence-based typing method: definition of the B*1520 allele. *Tissue Antigens 44:* 217-24, 1994.

Dormoy, A., Urlacher, A. and Tongio, M.M. Complexity of the HLA-DP region: RFLP analysis versus PLT typing and oligotyping. *Hum Immunol 34:* 39-46, 1992.

Dupont, B. Editorial. "Phototyping" for HLA: the beginning of the end of HLA typing as we know it. *Tissue Antigens 46:* 353-354, 1995.

duToit, E., Emmanuel, J.C., West, G., Taljaard, D.G. and Oudshoorn, M. HLA frequencies in black Zimbabweans. *Tissue Antigens 36:* 122-4, 1990.

duToit, E., MacGregor, K.J., Taljaard, D.G. and Oudshoorn, M. HLA-A, B, C, DR and DQ polymorphisms in three South African population groups: South African Negroes, Cape Coloureds and South African Caucasoids. *Tissue Antigens 31:* 109-25, 1988.

Edwards, P. A., C. M. Smith, A. M. Neville and M. J. O'Hare. A human-hybridoma system based on a fast-growing mutant of the ARH-77 plasma cell leukemia-derived line. *European Journal of Immunology* 12: 641-648: 1982.

Eiermann, T.H., Fakler, J. and Goldmann, S.F. The incidence of DPB1 differences between serological and mixed lymphocyte culture matched unrelated individuals: implications for selection of bone marrow donors. *Bone Marrow Transplant 9:* 157-60, 1992.

Engelhard, V. Structure of peptides associated with MHC class I molecules. *Curr Opin Immunol* 6(1): 13-23: 1994.

Erlich, H., Bugawan, T., Begovich, A.B., Scharf, S., Griffith, R., Saiki, R., Higuchi, R. and Walsh, P.S. HLA-DR, DQ and DP typing using PCR amplification and immobilized probes. *Eur J Immunogen 18:* 33-55, 1991a.

Erlich, H.A. and Gyllensten, U.B. Shared epitopes among HLA class II alleles: gene conversion, common ancestry and balancing selection. *Immunol Today 12:* 411-4, 1991b.

Falk, K., Rotzschke, O., and Rammensee, H.G. Cellular peptide composition governed by major histocompatibility complex class I molecules. *Nature* 348(6298): 248-51: 1990.

Falk, K., Rotzschke, O., Stevanovic, S., Jung, G. and Rammensee, H.G. Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. *Nature 351:* 290-6, 1991.

Fehling H.J., Swat W., Laplace C., Kuhn R., Rajewsky K., Muller U. and von Boehmer. H. MHC class I expression in mice lacking the proteasome subunit LMP-7. *Science* 265(5176): 1234-7: 1994.

Fernandez-Vina, M., Lazaro, A.M., Sun, Y., Miller, S., Forero, L. and Stastny, P. Population diversity of B-locus alleles observed by high resolution DNA typing. *Tissue Antigens 45:* 153-168, 1995.

Fernandez-Vina, M.A., Falco, M., Sun, Y. and Stastny, P. DNA typing for HLA class I alleles: I. Subsets of HLA-A2 and of -A28. *Hum Immunol 33:* 163-73, 1992.

Ferrone, S. and Pellegrino, M.A. HL-A antigens, antibody, and complement in the lymphocytotoxic reaction. *Contemp Top Mol Immunol 2:* 185-235, 1973.

Fruci, D., Rovero, P., Falasca, G., Chersi, A., Sorrentino, R., Butler, R., Tanigaki, N. and Tosi, R. Anchor residue motifs of HLA class-I-binding peptides analyzed by the direct binding of synthetic peptides to HLA class I alpha chains. *Hum Immunol 38:* 187-92, 1993.

Gallardo, D., Argüello, R., Rojas, I., Little, A. M., Avakian, H. and Madrigal, J. A. A high resolution HLA class I typing method with group specific amplification and novel SSO probe set. *Human Immunology* 49: 122-122, 1996.

Gallimore, A., Cranage, M., Cook, N., Almond, N., Bootman, J., Rud, E., Silvera, P., Dennis, M., Corcoran, T., Stott, J. and et, a.l. Early suppression of SIV replication by CD8+ nef-specific cytotoxic T cells in vaccinated macaques. *Nat Med 1:* 1167-73, 1995.

Gao, X., Jakobsen, I.B. and Serjeantson, S.W. Characterization of the HLA-A polymorphism by locus-specific polymerase chain reaction amplification and oligonucleotide hybridization. *Hum Immunol 41:* 267-79, 1994.

Garboczi, D. N., P. Ghosh, U. Utz, Q. R. Fan, W. E. Biddison and D. C. Wiley. Structure of the complex between human T-cell receptor, viral peptide and HLA-A2. *Nature* 384: 134-141: 1996.

Garcia, K. C., M. Degano, R. L. Stanfield, A. Brunmark, M. R. Jackson, P. A. Peterson, L. Teyton and I. A. Wilson. An αβ T Cell Receptor Structure at 2.5 Å and Its Orienation in the TCR-MHC Complex. *Science* 274: 209-219: 1996.

Garcia, P.J., Mantilla, P., Garcia, O.E. and Manzano, F.M. Routine HLA DRB/DQB oligonucleotide typing by a non-radioactive dot-blot micromethod. *J Immunol Methods 180:* 35-43, 1995.

Geraghty, D.E., Koller, B.H., Pei, J. and Hansen, J.A. Examination of four HLA class I pseudogenes. Common events in the evolution of HLA genes and pseudogenes. *J Immunol 149:* 1947-56, 1992.

Germain R.N., Dorf M.E and Benacerraf B. Inhibition of T-lymphocyte-mediated tumor-specific lysis by alloantisera directed against the H-2 serological specificities of the tumor. *J Exp Med* 142(4): 1023-8: 1975.

Giorda, R., Lampasona, V., Kocova, M. and Trucco, M. Non-radioisotopic typing of human leukocyte antigen class II genes on microplates. *Biotechniques 15:* 918-25, 1993.

Glynne R., Powis S., Beck S., Kelly A., Kerr L.A and Trowsdale J. A proteasome-related gene between the two ABC transporter loci in the class II region of the human MHC. *Nature* 353(6342): 357-60: 1991.

Gorga, J.C., Brown, J.H., Jardetzky, T., Wiley, D.C. and Strominger, J.L. Crystallization of HLA-DR antigens. *Res Immunol 142:* 401-7, 1991.

Gotch, F.M., Kelly, C., Ellis, S.A., Wallace, L., Rickinson, A.B., van der Poel, J., Crumpton, M.J. and McMichael, A.J. Characterization of the HLA-A2.2 subtype: T cell evidence for further heterogeneity. *Immunogenetics 21:* 11-23, 1985.

Gotch, F., McMichael, A. and Rothbard, J. Recognition of influenza A matrix protein by HLA-A2-restricted cytotoxic T lymphocytes. Use of analogues to orientate the matrix peptide in the HLA-A2 binding site. *J Exp Med 168:* 2045-57, 1988.

Gotch, F., McMichael, A., Smith, G. and Moss, B. Identification of viral molecules recognized by influenza-specific human cytotoxic T lymphocytes. *J Exp Med 165:* 408-16, 1987.

Gumperz, J.E., Litwin, V., Phillips, J.H., Lanier, L.L. and Parham, P. The Bw4 public epitope of HLA-B molecules confers reactivity with natural killer cell clones that express NKB1, a putative HLA receptor. *J Exp Med 181:* 1133-44, 1995.

Guttridge, M.G., Burr, C. and Klouda, P.T. Identification of HLA-B35, B53, B18, B5, B78, and B17 alleles by the polymerase chain reaction using sequence-specific primers (PCR-SSP). *Tissue Antigens 44:* 43-6, 1994.

Guttridge, M.G., Marsh, S.G. and Klouda, P.T. A new HLA-A2 variant (HLA-A2-SLU) identified by one-dimensional isoelectric focusing. *Tissue Antigens 39:* 32-4, 1992.

Halle, L., Mbayo, K., Lurhuma, Z., Salmon, D., Martageix, C., Castellano, F. and Kaplan, C. HLA-A, B, C, DR and DQ polymorphism in Zaireans. *Tissue Antigens 44:* 196-9, 1994

Helenius, A. How N-linked oligosaccharides affect glycoprotein folding in the endoplasmic reticulum. *Mol Biol Cell* 5(3): 253-65: 1994.

Higuchi, R., von Beroldingen, C., Sensabaugh, G.F. and Erlich, H.A. DNA typing from single hairs. *Nature 332:* 543-6, 1988.

Hildebrand, W.H., Domena, J.D., Shen, S.Y., Lau, M., Terasaki, P.I., Bunce, M., Marsh, S.G., Guttridge, M.G., Bias, W.B. and Parham, P. HLA-B15: a widespread and diverse family of HLA-B alleles. *Tissue Antigens 43:* 209-18, 1994a.

Hildebrand, W.H., Domena, J.D., Shen, S.Y., Marsh, S.G., Bunce, M., Guttridge, M.G., Darke, C. and Parham, P. The HLA-B7Qui antigen is encoded by a new subtype of HLA-B27 (B*2708). *Tissue Antigens 44:* 47-51, 1994b.

Hildebrand, W.H., Madrigal, J.A., Little, A.M. and Parham, P. HLA-Bw22: a family of molecules with identity to HLA-B7 in the alpha 1-helix. *J Immunol 148:* 1155-62, 1992.

Hill, A., Allsopp, C., Kwiatkowski, D., Anstey, N., Greenwood, B. and McMichael, A. HLA class I typing by PCR: HLA-B27 and an African B27 subtype [see comments]. *Lancet* 337(8742): 640-2: 1991a.

Hill, A.V., Allsopp, C.E., Kwiatkowski, D., Anstey, N.M., Twumasi, P., Rowe, P.A., Bennett, S., Brewster, D., McMichael, A.J. and Greenwood, B.M. Common west African HLA antigens are associated with protection from severe malaria [see comments]. *Nature 352:* 595-600, 1991b.

Hill, A.V., Elvin, J., Willis, A.C., Aidoo, M., Allsopp, C.E., Gotch, F.M., Gao, X.M., Takiguchi, M., Greenwood, B.M., Townsend, A.R. and et, a. Molecular analysis of the association of HLA-B53 and resistance to severe malaria [see comments]. *Nature 360:* 434-9, 1992.

Hilton, R., Lombardi, G. and Lechler, R. (1996) MHC molecules in transplantation and tolerance. In: HLA and MHC: genes, molecules and function (eds), p407-431 Bios Scientific, Oxford.

Hirschorn, K., Bach, F., Kolodny, R.L. and Firschen, I.L. Immune response and mitosis of human peripheral blood lymphocytes in vitro. *Science 142:* 1185-1187, 1963.

Hongming, F., Tilanus, M., M., v.E. and Giphart, M. Reduced complexity of RFLP for HLA-DR typing by the use of a DR beta 3' cDNA probe. *Tissue Antigens 28:* 129-35, 1986.

Howard, J. Supply and transport of peptides presented by class I MHC molecules. *Curr Opin Immunol* 7(1): 69-76: 1995.

Hviid, T.V., Madsen, H.O. and Morling, N. HLA-DPB1 typing with polymerase chain reaction and restriction fragment length polymorphism technique in Danes. *Tissue Antigens 40:* 140-4, 1992.

Hyldig, N.J., Odum, N., Morling, N. and Svejgaard, A. Restriction fragment polymorphism (RFLP) of a "new" HLA-DP specificity, CDP-HEI. *Tissue Antigens 31:* 161-3, 1988.

Imanishi, I., Akaza, T., Kimura, A., Tokunaga, K. and Gojobori, T. (1992a) Allele and haplotype frequencies for HLA and complement loci in various ethnic groups. In: HLA 1991. Proceedings of the Eleventh International Histocompatibility Workshop and Conference. (eds), Vol 1, p1065-220, Oxford University Press, Oxford.

Imanishi, T., Wakisaka, A. and Gojobori, T. (1992b) Genetic relationships among various human populations indicated by MHC polymorphism. In: HLA 1991. Proceedings of the Eleventh International Histocompatibility Workshop and Conference. (eds), Vol 1, p627-32 Oxford University Press, Oxford.

Inoko, H., Ando, A., Ito, M. and Tsuji, K. Southern hybridization analysis of DNA polymorphism in the HLA-D region. *Hum Immunol 16:* 304-13, 1986.

Ishikawa, Y., K. Tokunaga, H. Tanaka, M. Nishimura, M. Muraoka, Y. Fujii, T. Akaza, K. Tadokoro and T. Juji. HLA-A null allele with a stop codon, HLA-A*0215N, identified in a homozygous state in a healthy adult. *Immunogenetics* 43(1-2): 1-5: 1996.

Itescu, S., Rose, S., Dwyer, E. and Winchester, R. Grouping HLA-B locus serologic specificities according to shared structural motifs suggests that different peptide-anchoring pockets may have contrasting influences on the course of HIV-1 infection. *Hum Immunol 42:* 81-9, 1995.

Jaraquemada, D., Navarrete, C., Ollier, W., Awad, J., Okoye, R. and Festenstein, H. HLA-Dw specificity assignments are independent of HLA-DQ, HLA-DR, and other class II specificities and define a biologically important segregant series which strongly activates a functionally distinct T cell subset. *Hum Immunol 16:* 259-70, 1986.

Jones, E., P. Goodfellow, J. Bodmer and W. Bodmer. Serological identification of HL-A-linked human "Ia-type" antigens. *Nature* 256(5519): 651-2: 1975.

Kaminski, E.R., Hows, J.M., Bridge, J., Davey, N.J., Brookes, P.A., Green, J.E., Goldman, J.M. and Batchelor, J.R. Cytotoxic T lymphocyte precursor (CTL-p) frequency analysis in unrelated donor bone marrow transplantation: two case studies. *Bone Marrow Transplant 8:* 47-50, 1991.

Kast, W.M., Brandt, R.M., Sidney, J., Drijfhout, J.W., Kubo, R.T., Grey, H.M., Melief, C.J. and Sette, A. Role of HLA-A motifs in identification of potential CTL epitopes in human papillomavirus type 16 E6 and E7 proteins. *J Immunol 152:* 3904-12, 1994.

Kato, N., Kikuchi, A., Kano, K., Egawa, K. and Takiguchi, M. Molecular analysis of a novel HLA-A33 subtype associated with HLA-B44. *Tissue Antigens 41:* 211-213, 1993.

Kato, K., Trapani, J.A., Allopenna, J., Dupont, B. and Yang, S.Y. Molecular analysis of the serologically defined HLA-Aw19 Antigens. A genetically distinct family of HLA-A antigens comprising A29, A31, A32 and Aw33, but probably not A30. *Journal of Immunology 143:* 3371-3378, 1989.

Kawai, S., Maekawajiri, S., Tokunaga, K., Juji, T. and Yamane, A. A simple method of HLA-DRB typing using enzymatically amplified DNA and immobilized probes on microtiter plate. *Hum Immunol 41:* 121-6, 1994.

Kelly A., Powis S.H., Glynne R., Radley E., Beck S. and Trowsdale J. Second proteasome-related gene in the human MHC class II region. *Nature* 353(6345): 667-8: 1991.

Kimura, A. and Sasazuki, T. (1992) Eleventh International Histocompatibility Workshop reference protocol for the HLA DNA-typing technique. In: HLA 1991. Proceedings of the Eleventh International Histocompatibility Workshop and Conference. (eds), p397-419 Oxford University Press, Oxford.

Kissmeyer, N.F., Svejgaard, A. and Hauge, M. The HL-A system defined with lymphocytotoxic and platelet antibodies in relation to kidney transplantation. *Transplant Proc 1:* 357-61, 1969.

Kleijmeer M.J., Kelly A., Geuze H.J., Slot J.W. Townsend A. and Trowsdale J. Location of MHC-encoded transporters in the endoplasmic reticulum and cis-Golgi. *Nature* 357(6376): 342-4: 1992.

Knipper, A.J., Hinney, A., Schuch, B., Enczmann, J., Uhrberg, M. and Wernet, P. Selection of unrelated bone marrow donors by PCR-SSP typing and subsequent nonradioactive sequence-based typing for HLA DRB1/3/4/5, DQB1, and DPB1 alleles. *Tissue Antigens 44:* 275-84, 1994

Kovats, S., Main, E.K., Librach, C., Stubblebine, M., Fisher, S.J. and DeMars, R. A class I antigen, HLA-G, expressed in human trophoblasts. *Science 248:* 220-3, 1990.

Kraulis, P.J. MOLSCRIPT; a program to produce both detailed and schematic plots of protein structure. *Journal of Applied Crystallography 24:* 946-950, 1991.

Krausa, P., D. Barouch, J. G. Bodmer and M. J. Browning Rapid characterization of HLA Class I alleles by gene mapping using ARMS PCR. *European Journal of Immunogenetics 22:* 283-287, 1995d.

Krausa, P., Barouch, D., Bodmer, J.G., Hill, A.V.S., Mason, C., McMichael, A.J. and Browning, M.J. Characterisation of a novel HLA-A2 variant, A*0214, by ARMS-PCR and DNA sequencing. *Immunogenetics 41:* 50-50, 1995c

Krausa, P., Bodmer, J.G. and Browning, M.J. Defining the common subtypes of HLA A9, A10, A28 and A19 by use of ARMS/PCR. *Tissue Antigens 42:* 91-9, 1993

Krausa, P. and Browning, M.J. A comprehensive PCR-SSP typing system for identification of HLA-A locus alleles. *Tissue Antigens 47:* 237-244, 1996

Krausa, P., Brywka III, M., Savage, D., Hui, K.M., Bunce, M., Ngai, J.L.F., Teo, D.L.T., Ong, Y.W., Barouch, D., Allsop, C.E.M., Hill, A.V.S., McMichael, A.J., Bodmer, J.G. and Browning, M.J. Genetic polymorphism within HLA-A*02: significant allelic variation revealed in different population. *Tissue Antigens 45:* 223-231, 1995a.

Krausa, P., Carcassi, C., Orrú, S., Bodmer, J.G., Browning, M.J. and Contu, L. Defining the allelic variants of HLA-A30 in the Sardinian population using Amplification Refractory Mutation System-Polymerase Chain Reaction. *Human Immunology 44:* 35-44, 1995b.

Kubo, R.T., Sette, A., Grey, H.M., Appella, E., Sakaguchi, K., Zhu, N.Z., Arnott, D., Sherman, N., Shabanowitz, J., Michel, H. and et, a. Definition of specific peptide motifs for four major HLA-A alleles. *J Immunol 152:* 3913-24, 1994.

Kuhner, M.K. and Peterson, M.J. Genetic exchange in the evolution of the human MHC class II loci. *Tissue Antigens 39:* 209-15, 1992.

Latron, F., Moots, R., Rothbard, J.B., Garrett, T.P., Strominger, J.L. and McMichael, A. Positioning of a peptide in the cleft of HLA-A2 by complementing amino acid changes. *Proc Natl Acad Sci U S A 88:* 11325-9, 1991.

Lawlor, D.A., Warren, E., Taylor, P. and Parham, P. Gorilla class I major histocompatibility complex alleles: comparison to human and chimpanzee class I. *J Exp Med 174:* 1491-509, 1991.

Lawlor, D.A., Zemmour, J., Ennis, P.D. and Parham, P. Evolution of class-I MHC genes and proteins: from natural selection to thymic selection. *Annu Rev Immunol 8:* 23-63, 1990.

Lee, H.C., Ladd, C., Bourke, M.T., Pagliaro, E. and Tirnady, F. DNA typing in forensic science. I. Theory and background. *Am J Forensic Med Pathol 15:* 269-82, 1994.

Levine, J.E. and Yang, S.Y. SSOP typing of the Tenth International Histocompatibility Workshop reference cell lines for HLA-C alleles. *Tissue Antigens 44:* 174-83, 1994.

Lienert, K., G. Russ, S. Lester, G. Bennett, X. Gao and J. McCluskey. Stable inheritance of an HLA-"blank" phenotype associated with a structural mutation in the HLA-A*0301 gene. *Tissue Antigens* 48: 187-191: 1996.

Lin, X.Y., Wang, Y., Sun, J., Wang, F.Q., Ni, L.Q., Chang, L.Y. and Chen, S.S. Association of RFLP of HLA class I genes with Chinese ankylosing spondylitis patients. *Tissue Antigens 34:* 279-83, 1989.

Loiseau, P., al, D.R., David, F. and Colombani, J. Molecular HLA-class II typing: advantages and clinical application. *Nouv Rev Fr Hematol 33:* 461-3, 1991.

Lopez de Castro, J.A. HLA-B27 and HLA-A2 subtypes: structure, evolution and function. *Immunology today 10:* 239-246, 1989.

Maeda, M., Murayama, N., Ishii, H., Uryu, N., Ota, M., Tsuji, K. and Inoko, H. A simple and rapid method for HLA-DQA1 genotyping by digestion of PCR-amplified DNA with allele specific restriction endonucleases. *Tissue Antigens 34:* 290-8, 1989.

Madden, D.R., Garboczi, D.N. and Wiley, D.C. The antigenic identity of peptide-MHC complexes: a comparison of the conformations of five viral peptides presented by HLA-A2. *Cell 75:* 693-708, 1993.

Marsh, S.G.E. Nomenclature for factors of the HLA system, update April 1996. *Tissue Antigens 48:* 69-70, 1996.

Marsh, S.G. and Bodmer, J.G. HLA-DR and -DQ epitopes and monoclonal antibody specificity. *Immunol Today 10:* 305-12, 1989.

Marsh, S.G.E. and Bodmer, J.G. HLA Class II region nucleotide sequences 1995. *Tissue Antigens 46:* 258-280, 1995.

Marsh, S.G.E., Fauchet, R.M., Yang, E.K., Moses, J.R., Semana, G., Mervart, H., Chadanayingyong, D., Parham, P., Tongio, M.M., Patel, H., Guttridge, M., Kunmartinin, S., Chiewslip, P., Sujiachato, K., Tsuneto, L.T., Trabace, S. and Bodmer, J.G. (1992) Antigen Society no. 102: HLA-A2, -A68, -A9, -A23, -A24. In: HLA 1991. Proceedings of the Eleventh International Histocompatibility Workshop and Conference. (eds), Vol 1, p285-9, Oxford University Press, Oxford.

Matsui, M., Moots, R.J., McMichael, A.J. and Frelinger, J.A. Significance of the six peptide-binding pockets of HLA-A2.1 in influenza A matrix peptide-specific cytotoxic T-lymphocyte reactivity. *Hum Immunol 41:* 160-6, 1994.

Matsui, M., Moots, R.J., Warburton, R.J., Peace, B.A., Tussey, L.G., Quinn, D.G., McMichael, A.J. and Frelinger, J.A. Genetic evidence for difference between intracellular and extracellular peptides in influenza A matrix peptide-specific CTL recognition. *J Immunol 154:* 1088-96, 1995.

Matsumura, M., Fremont, D.H., Peterson, P.A. and Wilson, I.A. Emerging principles for the recognition of peptide antigens by MHC class I molecules [see comments]. *Science 257:* 927-34, 1992.

Mazzoleni, O., M. Pistillo, M. Falco, P. Tazzari and G. Ferrara. Characterization of two human monoclonal antibodies recognizing HLA-A30 and HLA-A3 + A31, respectively. *Tissue Antigens* 38(5): 224-7: 1991.

McDaniel, D.O., Acton, R.T., Barger, B.O., Koopman, W.J. and Reveille, J.D. Association of a 9.2-kilobase Pvu II class I major histocompatibility complex restriction fragment length polymorphism with ankylosing spondylitis. *Arthritis Rheum 30:* 894-900, 1987.

McMichael, A.J., Gotch, F.M., Santos-Aguado, J. and Strominger, J.L. Effect of mutations and variations of HLA-A2 on recognition of a virus peptide epitope by cytotoxic T lymphocytes. *Proc Natl Acad Sci U S A 85:* 9194-8, 1988.

McMichael, A.J. and Walker, B. Cytotoxic T lymphocyte epitopes: implications for HIV vaccines. *AIDS 8(suppl 1):* S155-S173, 1994.

Medintz, I., Chiriboga, L., McCurdy, L. and Kobilinsky, L. Restriction fragment length polymorphism and polymerase chain reaction-HLA DQ alpha analysis of casework urine specimens. *J Forensic Sci 39:* 1372-80, 1994.

Mempel, W., Albert, E. and Burger, A. Further evidence for a separate MLC-locus. *Tissue Antigens 2:* 250-4, 1972.

Mempel, W., Grosse, W.H., Albert, E. and Thierfelder, S. Atypical MLC reactions in HL-A-typed related and unrelated pairs. *Transplant Proc 5:* 401-8, 1973a.

Mempel, W., Grosse, W.H., Baumann, P., Netzel, B., Steinbauer, R.I., Scholz, S., Bertrams, J. and Albert, E.D. Population genetics of the MLC response: typing for MLC determinants using homozygous and heterozygous reference cells. *Transplant Proc 5:* 1529-34, 1973b.

Merel, P., Dupin, B., Comeau, F. and Vezon, G. Automation progress in PCR/SSP protocols: from DNA extraction to PCR setup (abstract). *Human Immunology* 41(supplement 1): 119-119: 1996.

Merritt, E.A. and Murphy, M.E.P. Raster3D version 2.0Å program for photorealistic molecular graphics. *Acta Cryst. D50:* 869-873, 1994.

Middleton, D., Williams, F., Cullen, C. and Mallon, E. Modification of an HLA-B PCR-SSOP typing system leading to improved allele determination. *Tissue Antigens 45:* 232-236, 1995.

Miller J.F., Vadas M.A., Whitelaw A. and Gamble J. Role of major histocompatibility complex gene products in delayed-type hypersensitivity. *Proc Natl Acad Sci U S A* 73(7): 2486-90: 1976.

Miller, S.A., Dykes, D.D. and Polesky, H.F. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res 16:* 1215, 1988.

Moller, E., Carlsson, B. and Wallin, J. Implication of structural class II gene polymorphism for the concept of serologic specificities. *Immunol Rev 85:* 107-28, 1985.

Morishima, Y., Kodera, Y., Hirabayashi, N., Tanimoto, M., Matsuyama, T., Horibe, K., Ohta, H., Mizuno, S., Morishita, Y., Yamauchi, T. and et, a.l. Low incidence of acute GVHD in patients transplanted with marrow from HLA-A,B,DR-compatible unrelated donors among Japanese. *Bone Marrow Transplant 15:* 235-9, 1995.

Morling, N., Jakobsen, B.K., Platz, P., Ryder, L.P., Svejgaard, A. and Thomsen, M. A "new" primed lymphocyte typing (PLT) defined DP-antigen associated with a private HLA--DR antigen. *Tissue Antigens 16:* 95-104, 1980.

Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G. and Erlich, H. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harb Symp Quant Biol 1:* 263-73, 1986.

Mullis, K.B. and Faloona, F.A. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods Enzymol 155:* 335-50, 1987.

Niedermann G., Butz S., Ihlenfeldt H.G., Grimm R., Lucchiari M., Hoschutzky H., Jung G. Maier B. and Eichmann K. Contribution of proteasome-mediated proteolysis to the hierarchy of epitopes presented by major histocompatibility complex class I molecules. *Immunity* 2(3): 289-99: 1995.

Newton, C.R., Graham, A., Heptinstall, L.E., Powell, S.J., Summers, C., Kalsheker, N., Smith, J.C. and Markham, A.F. Analysis of any point mutation in DNA. The amplification refractory mutation system (ARMS). *Nucleic Acids Res 17:* 2503-16, 1989.

Oh, S.H., Fleischhauer, K. and Yang, S.Y. Isoelectric focusing subtypes of HLA-A can be defined by oligonucleotide typing. *Tissue Antigens 41:* 135-42, 1993.

Oka, T., Matsunaga, H., Tokunaga, K., Mitsunaga, S., Juji, T. and Yamane, A. A simple method for detecting single base substitutions and its application to HLA-DPB1 typing. *Nucleic Acids Res 22:* 1541-7, 1994.

Olerup, O., Aldener, A. and Fogdell, A. HLA-DQB1 and -DQA1 typing by PCR amplification with sequence-specific primers (PCR-SSP) in 2 hours. *Tissue Antigens 41:* 119-34, 1993.

Olerup, O., Daniels, T. and Baxter, L.L. Correct sequence of the A*3001 allele obtained by PCR-SSP typing and automated nucleotide sequencing. *Tissue Antigens 44:* 265-7, 1994.

Olerup, O. and Zetterquist, H. HLA-DRB1*01 subtyping by allele-specific PCR amplification: a sensitive, specific and rapid technique. *Tissue Antigens 37:* 197-204, 1991.

Olerup, O. and Zetterquist, H. HLA-DR typing by PCR amplification with sequence-specific primers (PCR-SSP) in 2 hours: an alternative to serological DR typing in clinical practice including donor-recipient matching in cadaveric transplantation [see comments]. *Tissue Antigens 39:* 225-35, 1992.

Ortmann, B., M. Androlewicz and P. Cresswell. MHC class I/beta 2-microglobulin complexes associate with TAP transporters before peptide binding. *Nature* 368(6474): 864-7: 1994.

Payne, R., M. Tripp, J. Weigle, W. F. Bodmer and J. Bodmer. A new leukocyte isoantigen system in man. *Cold Spring Harbor Symp Quant Biol* 29: 285-295: 1964.

Parham, P. and Brodsky, F.M. Partial purification and some properties of BB7.2: A cytotoxic monoclonal antibody with specificity for HLA-A2 and a variant of HLA-A28. *Human Immunology 3:* 277-299, 1981.

Parham, P., Lomen, C.E., Lawlor, D.A., Ways, J.P., Holmes, N., Coppin, H.L., Salter, R.D., Wan, A.M. and Ennis, P.D. Nature of polymorphism in HLA-A, -B, and -C molecules. *Proc Natl Acad Sci U S A 85:* 4005-9, 1988.

Parham, P. Pictures of MHC restriction. *Nature* 384: 109-110: 1996.

Peters J.M. Proteasomes: protein degradation machines of the cell. *Trends Biochem Sci* 19(9): 377-82: 1994.

Piazza, A., Mayr, W.R., Contu, L., Amoroso, A., Borelli, I., Curtoni, E.S., Marcello, C., Moroni, A., Olivetti, E., Richiardi, P. and et, a.l. Genetic and population structure of four Sardinian villages. *Ann Hum Genet* 1985.

Plourde, P.J., Plummer, F.A., Pepin, J., Agoki, E., Moss, G., Ombette, J., Ronald, A.R., Cheang, M., D'Costa, L. and Ndinya, A.J. Human immunodeficiency virus type 1 infection in women attending a sexually transmitted diseases clinic in Kenya [see comments]. *J Infect Dis 166:* 86-92, 1992.

Powis, S., E. Deverson, W. Coadwell, A. Ciruela, N. Huskisson, H. Smith, G. Butcher and J. Howard. Effect of polymorphism of an MHC-linked transporter on the peptides assembled in a class I molecule. *Nature* 357(6375): 211-5: 1992.

Rammensee, H.G., Friede, T. and Stevanoviic, S. MHC ligands and peptide motifs: first listing. *Immunogenetics 41:* 178-228, 1995.

Rinke, d.W.T., Vloemans, S., van den Elsen, P., Haworth, A. and Stern, P.L. Differential expression of the HLA class I multigene family by human embryonal carcinoma and choriocarcinoma cell lines. *J Immunol 144:* 1080-7, 1990.

Rosa, F., Fellous, M., Dron, M., Tovey, M. and Revel, M. Presence of an abnormal beta 2-microglobulin mRNA in Daudi cells: induction by interferon. *Immunogenetics 17:* 125-31, 1983.

Rosen-Bronson, S., Wagner, A.G., Stewart, D., Herbert, S. and Hurley, C.K. DNA sequencing of a new HLA-A allele using locus specific PCR amplification. *Human Immunology 34:* 15, 1992.

Rotzschke O., Falk K., Deres K., Schild H., Norda M., Metzger J., Jung G. and Rammensee H.G. Isolation and analysis of naturally processed viral peptides as recognized by cytotoxic T cells [see comments]. *Nature* 348(6298): 252-4: 1990.

Rotzschke, O., Falk, K., Stevanovic, S., Jung, G. and Rammensee, H.G. Peptide motifs of closely related HLA class I molecules encompass substantial differences. *Eur J Immunol 22:* 2453-6, 1992.

Rowland-Jones, S., Sutton, J., Ariyoshi, K., Dong, T., Gotch, F., McAdam, S., Whitby, D., Sabally, S., Gallimore, A., Corrah, T. and et, a.l. HIV-specific cytotoxic T-cells in HIV-exposed but uninfected Gambian women. *Nat Med 1:* 59-64, 1995.

Rowland-Jones, S.L., Nixon, D.F., Aldhous, M.C., Gotch, F., Ariyoshi, K., Hallam, N., Kroll, J.S., Froebel, K. and McMichael, A. HIV-specific cytotoxic T-cell activity in an HIV-exposed but uninfected infant. *Lancet 341:* 860-1, 1993.

Rozemuller, E.H., Bouwens, A.G., Bast, B.E. and Tilanus, M.G. Assignment of HLA-DPB alleles by computerized matching based upon sequence data. *Hum Immunol 37:* 207-12, 1993.

Rubocki, R.J., Wisecarver, J.L., Hook, D.D., Cox, S.M. and Beisel, K.W. Histocompatibility screening by molecular techniques: use of polymerase chain reaction products and heteroduplex formation. *J Clin Lab Anal 6:* 337-41, 1992.

Sadler, A.M., Petronzelli, F., Krausa, P., Marsh, S.G., Guttridge, M.G., Browning, M.J. and Bodmer, J.G. Low-resolution DNA typing for HLA-B using sequence-specific primers in allele- or group-specific ARMS/PCR. *Tissue Antigens 44:* 148-54, 1994.

Saiki, R.K., Bugawan, T.L., Horn, G.T., Mullis, K.B. and Erlich, H.A. Analysis of enzymatically amplified beta-globin and HLA-DQ alpha DNA with allele-specific oligonucleotide probes. *Nature 324:* 163-6, 1986.

Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A. and Arnheim, N. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science 230:* 1350-4, 1985.

Saiki, R.K., Walsh, P.S., Levenson, C.H. and Erlich, H.A. Genetic analysis of amplified DNA with immobilized sequence-specific oligonucleotide probes. *Proc Natl Acad Sci U S A 86:* 6230-4, 1989.

Salazar, M., Yunis, J.J., Delgado, M.B., Bing, D. and Yunis, E.J. HLA-DQB1 allele typing by a new PCR-RFLP method: correlation with a PCR-SSO method. *Tissue Antigens 40:* 116-23, 1992.

Sanger, F., Nicklen, S. and Coulson, A.R. DNA sequencing with chain-terminating inhibitors. *Proceedings National Academy Sciences USA 74:* 5463-5467, 1977.

Santamaria, P., Boyce, J.M., Lindstrom, A.L., Barbosa, J.J., Faras, A.J. and Rich, S.S. HLA class II "typing": direct sequencing of DRB, DQB, and DQA genes. *Hum Immunol 33:* 69-81, 1992.

Santamaria, P., Lindstrom, A.L., Boyce, J.M., Myster, S.H., Barbosa, J.J., Faras, A.J. and Rich, S.S. HLA class I sequence-based typing. *Hum Immunol 37:* 39-50, 1993.

Saper, M.A., Bjorkman, P.J. and Wiley, D.C. Refined structure of the human histocompatibility antigen HLA-A2 at 2.6 A resolution. *J Mol Biol 219:* 277-319, 1991.

Satsangi, J., Jewell, D.P., Welsh, K., Bunce, M. and Bell, J.I. Effect of heparin on polymerase chain reaction [letter]. *Lancet 343:* 1509-10, 1994.

Schrader J.W., and Edelman G.M. Participation of the H-2 antigens of tumor cells in their lysis by syngeneic T cells. *J Exp Med* 143(3): 601-14: 1976.

Scott, J. and J. Dawson. MHC class I expression and transport in a calnexin-deficient cell line. *J Immunol* 155(1): 143-8: 1995.

Segurado, O.G., Iglesias, C.P., Vicario, J.L., Corell, A., Regueiro, J.R. and Arnaiz, V.A. Shared SstI RFLPs by HLA-Aw19, A23/24 and A3/11 crossreacting groups. *Tissue Antigens 35:* 206-10, 1990.

Selvakumar, A., Granja, C.B., Salazar, M., Alosco, S.M., Yunis, E.J. and Dupont, B. A novel subtype of A2 (A*0217) isolated from the South American Indian B-cell line AMALA. *Tissue Antigens 45:* 343-347, 1995.

Sheehy, M.J., Sondel, P.M., Bach, M.L., Wank, R. and Bach, F.H. HL-A LD (lymphocyte defined) typing: a rapid assay with primed lymphocytes. *Science 188:* 1308-10, 1975.

Sheldon, E.L., Kellogg, D.E., Watson, R., Levenson, C.H. and Erlich, H.A. Use of nonisotopic M13 probes for genetic analysis: application to HLA class II loci. *Proc Natl Acad Sci U S A 83:* 9085-9, 1986.

Shintaku, S., Fukuda, Y., Kimura, A., Hoshino, S., Tashiro, H., Sasazuki, T. and Dohi, K. DNA conformation polymorphism analysis of DR52 associated HLA-DR antigens by polymerase chain reaction: a simple, economical and rapid examination for HLA matching in transplantation. *Jpn J Med Sci Biol 46:* 165-81, 1993.

Sorrentino, R., Cascino, I. and Tosi, R. Subgrouping of DR4 alleles by DNA heteroduplex analysis. *Hum Immunol 33:* 18-23, 1992.

Sorrentino, R., Potolicchio, I., D'Amato, M. and Tosi, R. The HLA DP locus in bone marrow transplantation. Probeless genomic typing of the DPB1 alleles. *Bone Marrow Transplant 1:* 17-9, 1993.

Sorrentino, R., Potolicchio, I., Ferrara, G.B. and Tosi, R. A new approach to HLA-DPB1 typing combining DNA heteroduplex analysis with allele-specific amplification and enzyme restriction. *Immunogenetics 36:* 248-54, 1992.

Spurkland, A., Knutsen, I., Markussen, G., Vartdal, F., Egeland, T. and Thorsby, E. HLA matching of unrelated bone marrow transplant pairs: direct sequencing of in vitro amplified HLA-DRB1 and -DQB1 genes using magnetic beads as solid support. *Tissue Antigens 41:* 155-64, 1993.

Srivastava, P., H. Udono, N. Blachere and Z. Li. Heat shock proteins transfer peptides during antigen processing and CTL priming. *Immunogenetics* 39(2): 93-8: 1994.

Starling, G.C., Witkowski, J.A., Speerbrecher, L.S., McKinney, S.K., Hansen, J.A. and Choo, S.Y. A novel HLA-A*8001 allele identified in the African-American population. *Human Immunology 39:* 163-168, 1994.

Strauss, W.M. (1990) Preparation of Genomic DNA from Mammalian Tissue. In: Current Protocols in Molecular Biology (eds), p2.2.1-2.2.3 Wiley, New York.

Storkus, W. J., D. N. Howell, R. D. Salter, J. R. Dawson and P. Cresswell. NK susceptibility varies inversely with target cell class I HLA antigen expression. *Journal of Immunology* 138: 1657-1659: 1987.

Sudo, T., Kamikawaji, N., Kimura, A., Date, Y., Savoie, C.J., Nakashima, H., Furuichi, E., Kuhara, S. and Sasazuki, T. Differences in MHC Class I self peptide repertoires among HLA-A2 subtypes. *Journal of Immunology 155:* 4749-4756, 1995.

Suh, W., M. Cohen-Doyle, K. Fruh, K. Wang, P. Peterson and D. Williams. Interaction of MHC class I molecules with the transporter associated with antigen processing. *Science* 264(5163): 1322-6: 1994.

Sutton, J., Rowland, J.S., Rosenberg, W., Nixon, D., Gotch, F., Gao, X.M., Murray, N., Spoonas, A., Driscoll, P., Smith, M. and et, a.l. A sequence pattern for peptides presented to cytotoxic T lymphocytes by HLA B8 revealed by analysis of epitopes and eluted peptides. *Eur J Immunol 23:* 447-53, 1993.

Tanaka, M., Abe, J., Kohsaka, T. and Tanae, A. Analysis of HLA-DQA1 in Japanese patients with type 1 diabetes mellitus, using DNA-PCR-RFLP typing. *Acta Paediatr Jpn 34:* 46-51, 1992.

Tanigaki, N., Fruci, D., Chersi, A., Falasca, G., Tosi, R. and Butler, R.H. HLA-A2-binding peptides cross-react not only within the A2 subgroup but also with other HLA-A-locus allelic products. *Hum Immunol 39:* 155-62, 1994.

Terasaki, P.I. and McClelland, J.D. Microdoplet assay of human serum cytotoxins. *Nature 204:* 998-1007, 1964.

Thein, S.L. and Wallace, R.B. (1986) The use of synthetic oligonucleotides as specific hybridisation probes in the diagnosis of genetic disorders. In: Human genetic disorders: a practical approach. (eds), p35-50 IRL Press, Herndon, Virginia.

Tiercy, J.M., Djavad, N., Rufer, N., Speiser, D.E., Jeannet, M. and Roosnek, E. Oligotyping of HLA-A2, -A3, and -B44 subtypes. Detection of subtype incompatibilities between patients and their serologically matched unrelated bone marrow donors. *Hum Immunol 41:* 207-15, 1994.

Tiercy, J.M., Morel, C., Freidel, A.C., Zwahlen, F., Gebuhrer, L., Betuel, H., Jeannet, M. and Mach, B. Selection of unrelated donors for bone marrow transplantation is improved by HLA class II genotyping with oligonucleotide hybridization. *Proc Natl Acad Sci U S A 88:* 7121-5, 1991.

Tiercy, J.M., Zwahlen, F., Jeannet, M. and Mach, B. [Current approach to HLA typing for bone marrow transplantation: oligonucleotide typing by hybridization on DNA amplified by polymerase chain reaction]. *Schweiz Med Wochenschr 119:* 1344-6, 1989.

Tilanus, M.G., Hongming, F., van Eggermond, M., v d Bijl, M., D'Amaro, J., Schreuder, G.M., de Vries, R. and Giphart, M.J. An overview of the restriction fragment length polymorphism of the HLA-D region: its application to individual D-, DR- typing by computerized analyses. *Tissue Antigens 28:* 218-27, 1986.

Townsend, A., Elliott, T., Cerundolo, V., Foster, L., Barber, B. and Tse, A. Assembly of MHC class I molecules analyzed in vitro. *Cell 62:* 285-95, 1990.

Townsend, A. R., Gotch F.M. and Davey J. Cytotoxic T cells recognize fragments of the influenza nucleoprotein. *Cell* 42: 457-467: 1985.

Townsend, A. R., McMichael A. J., Carter N. P., Huddleston J. A. and Brownlee G. G. Cytotoxic T cell recognition of the influenza nucleoprotein and hemagglutinin expressed in transfected mouse L cells. *Cell* 39(1): 13-25: 1984.

Townsend, A., C. Ohlen, J. Bastin, H. G. Ljunggren, L. Foster and K. Karre. Association of class I major histocompatibility heavy and light chains induced by viral peptides. *Nature* 340(6233): 443-8: 1989.

Townsend, A. R., Rothbard J., Gotch F.M., Bahadur G., Wraith D. and McMichael A.J. The epitopes of influenza nucleoprotein recognized by cytotoxic T lymphocytes can be defined with short synthetic peptides. *Cell* 44(6): 959-68: 1986.

Townsend, A. and J. Trowsdale. The transporters associated with antigen presentation. *Semin Cell Biol* 4(1): 53-61: 1993.

Trowsdale J., Hanson I., Mockridge I., Beck S., Townsend A. and Kelly A. Sequences encoded in the class II region of the MHC related to the 'ABC' superfamily of transporters [see comments]. *Nature* 348(6303): 741-4: 1990.

Tussey, L.G., Matsui, M., Rowland-Jones, S., Warburton, R., Frelinger, J.A. and McMichael, A. Analysis of mutant HLA-A2 molecules. Differential effects on peptide binding and CTL recognition. *J Immunol 152:* 1213-21, 1994

Uhrberg, M., Hinney, A., Enczmann, J. and Wernet, P. Analysis of the HLA-DR gene locus by temperature gradient gel electrophoresis and its application for the rapid selection of unrelated bone marrow donors. *Electrophoresis 15:* 1044-50, 1994.

Unanue E.R., Beller D.I., Lu C.Y. and Allen P.M. Antigen presentation: comments on its regulation and mechanism. *J Immunol* 132(1): 1-5: 1984.

Utz, U., Koenig, S., Coligan, J.E. and Biddison, W.E. Presentation of three different viral peptides, HTLV-1 Tax, HCMV gB, and influenza virus M1, is determined by common structural features of the HLA-A2.1 molecule. *J Immunol 149:* 214-21, 1992.

van der Poel, J., Pool, J., Goulmy, E., Giphart, M.J. and van Rood, J. Recognition of distinct epitopes on the HLA-A2 antigen by cytotoxic T lymphocytes. *Hum Immunol 16:* 247-58, 1986.

van Leeuwen, A., Schuit, H.R. and van Rood, J. Typing for MLC (LD). II. The selection of nonstimulator cells by MLC inhibition tests using SD-indentical stimulator cells (MISIS) and fluorescence antibody studies. *Transplant Proc 5:* 1539-42, 1973.

van Leeuwen, A., Termijtelen, A., Shaw, S. and van Rood, J. Recognition of a polymorphic monocyte antigen in HLA. *Nature 298:* 565-7, 1982.

van Rood, J. J. and A. van Leeuwen. Leucocyte grouping. A method and its application. *Journal of Clinical Investigation* 42: 1382: 1963.

Vaughan, R.W. PCR-SSO typing for HLA-DRB alleles. *Eur J Immunogen 18:* 69-80, 1991.

Versluis, L.F., Rozemuller, E., Tonks, S., Marsh, S.G., Bouwens, A.G., Bodmer, J.G. and Tilanus, M.G. High-resolution HLA-DPB typing based upon computerized analysis of data obtained by fluorescent sequencing of the amplified polymorphic exon 2. *Hum Immunol 38:* 277-83, 1993.

Wells, R.S. and Parham, P. (1996) HLA Class I genes - structure and diversity. In: HLA and MHC: genes, molecules and function (eds), p77-94 Bios Scientific, Oxford.

Williams, D. The Merck Frosst Award Lecture 1994/La conference Merck Frosst 1994. Calnexin: a molecular chaperone with a taste for carbohydrate. *Biochem Cell Biol* 73(3-4): 123-32: 1995.

Williams, D.B., Vassilakos, A. and Suh, W.-K. Peptide presentation by MHC class I molecules. *Trends in Cell Biology 6:* 267-273, 1996.

Wordsworth, P. PCR-SSO typing in HLA-disease association studies. *Eur J Immunogen 18:* 139-46, 1991.

Yang, S.Y. (1989a) Population analysis of Class I HLA antigens by one-dimensional isoelectric gel electrophoresis: Workshop summary report. In: Immunobiology of HLA. Histocompatibility testing 1987. (eds), Vol 1, p309-31 Springer-Verlag, New York.

Yang, S.Y. (1989b) A standardised method for detection of HLA-A and HLA-B alleles by one-dimensional isoelectric focusing (IEF) gel electrophoresis. In: Immunobiology of HLA. Histocompatibility testing 1987. (eds), Vol 1, p332-5 Springer-Verlag, New York.

Yoshida, M., Kimura, A., Numano, F. and Sasazuki, T. Polymerase-chain-reaction-based analysis of polymorphism in the HLA-B gene. *Hum Immunol 34:* 257-66, 1992.

Yunis, E.J. and Amos, D.B. Three closely linked genetic systems relevant to transplantation. *Proc Natl Acad Sci U S A 68:* 3031-5, 1971.

Yunis, I., Salazar, M. and Yunis, E.J. HLA-DR generic typing by AFLP. *Tissue Antigens 38:* 78-88, 1991.

Zemmour, J., Gumperz, J.E., Hildebrand, W.H., Ward, F.E., Marsh, S.G., Williams, R.C. and Parham, P. The molecular basis for reactivity of anti-Cw1 and anti-Cw3 alloantisera with HLA-B46 haplotypes. *Tissue Antigens 39:* 249-57, 1992a.

Zemmour, J., Little, A.M., Schendel, D. J. and Parham, P. The HLA-A, B "negative" mutant cell line C1R expresses a novel HLA-B35 allele, which also has a point mutation in the translation initiation codon. *Journal of Immunology* 148: 1941-1948: 1992b.

Zemmour, J. and Parham, P. HLA class I nucleotide sequences, 1991. *Hum Immunol 31:* 195-206, 1991.

Zemmour, J. and Parham, P. Class I nucleotide sequences, 1992. *Tissue Antigens 40:* 221-228, 1992.

Zetterquist, H. and Olerup, O. Identification of the HLA-DRB1*04, -DRB1*07, and -DRB1*09 alleles by PCR amplification with sequence-specific primers (PCR-SSP) in 2 hours. *Hum Immunol 34:* 64-74, 1992.

Zhang, L., Li, S.G., Vandekerckhove, B., Termijtelen, A., van Rood, J. and Claas, F.H. Analysis of cytotoxic T cell precursor frequencies directed against individual HLA-A and -B alloantigens. *J Immunol Methods 121:* 39-45, 1989.

Zinkernagel, R. Restriction by H-2 gene complex of transfer of cell-mediated immunity to Listeria monocytogenes. *Nature* 251(5472): 230-3: 1974.

Zinkernagel, R. H-2 restriction of virus-specific cytotoxicity across the H-2 barrier. Separate effector T-cell specificities are associated with self-H-2 and with the tolerated allogeneic H-2 in chimeras. *J Exp Med* 144(4): 933-45: 1976.

Zinkernagel, R. and P. Doherty. Immunological surveillance against altered self components by sensitised T lymphocytes in lymphocytic choriomeningitis. *Nature* 251(5475): 547-8: 1974a.

Zinkernagel, R. and P. Doherty. Restriction of in vitro T cell-mediated cytotoxicity in lymphocytic choriomeningitis within a syngeneic or semiallogeneic system. *Nature* 248(450): 701-2: 1974b.

Zinkernagel R.M. and Rosenthal K.L. Experiments and speculation on antiviral specificity of T and B cells. *Immunol Rev* 58: 131-55: 1981.

Publications resulting from this thesis.

Barouch, D*., Krausa, P*., Bodmer, J., Browning, M. J. and McMichael, A. J. Identification of a novel HLA-A2 subtype, HLA-A*0216. *Immunogenetics* 41(6): 388-388: 1995.

Browning, M. J., Krausa, P., Rowan, A., Bicknell, D. C., Bodmer, J. G. and Bodmer, W. F. Tissue typing the HLA-A locus from genomic DNA by sequence-specific PCR: comparison of HLA genotype and surface expression on colorectal tumor cell lines. *Proc Natl Acad Sci U S A* 90(7): 2842-5: 1993.

Carcassi, C., Krausa, P., Bodmer, J., Contu, L. and Browning, M. Characterization of HLA-A*02 subtypes in the Sardinian population. *Tissue Antigens* 46(5): 391-393: 1995.

Krausa, P., Barouch, D., Bodmer, J. G. and Browning, M. J. Rapid characterization of HLA Class I alleles by gene mapping using ARMS PCR. *European Journal of Immunogenetics* 22: 283-287: 1995.

Krausa, P., Barouch, D., Bodmer, J. G., Hill, A. V. S., Mason, C., McMichael, A. J. and Browning, M. J. Characterisation of a novel HLA-A2 variant, A*0214, by ARMS-PCR and DNA sequencing. *Immunogenetics* 41: 50-50: 1995.

Krausa, P., Bodmer, J. G. and Browning, M. J. Defining the common subtypes of HLA A9, A10, A28 and A19 by use of ARMS/PCR. *Tissue Antigens* 42(2): 91-9: 1993.

Krausa, P. and Browning, M. J. A comprehensive PCR-SSP typing system for identification of HLA-A locus alleles. *Tissue Antigens* 47: 237-244: 1996.

Krausa, P. and Browning, M.J. HLA-A2 Polymorphism and Immune Function. *European Journal of Immunogenetics* 23: 261-74: 1996.

Krausa, P., Browning, M. J., Moses, J. H. and Bodmer, J. G. HLA-A locus typing of the 4AOH cell panel by ARMS PCR. *European Journal of Immunugenetics* 21: 199-204: 1994.

Krausa, P., Brywka III, M., et al. Genetic polymorphism within HLA-A*02: significant allelic variation revealed in different population. *Tissue Antigens* 45(4): 223-231: 1995.

Krausa, P., Carcassi, C., Orrú, S., Bodmer, J. G., Browning, M. J. and Contu, L. Defining the allelic variants of HLA-A30 in the Sardinian population using Amplification Refractory Mutation System-Polymerase Chain Reaction. *Human Immunology* 44: 35-44: 1995.

Krausa, P., Moses, J., Bodmer, W., Bodmer, J. and Browning, M. HLA-A locus alleles identified by sequence specific PCR [letter]. *Lancet* 341(8837): 121-2: 1993.

Krausa, P., Young, D.M., Gotch, F. Identification of a new allele (A*2502) by PCR-SSP, *Immunogenetics* 45; 84: 1996.

* denotes an equal contribution to the publication.

Collaborations to which this work contributed.

Aidoo, M., Lalvani, A., Allsopp, C. E., Plebanski, M., Meisner, S. J., Krausa, P., Browning, M., Morris, J. S., Gotch, F., Fidock, D. A. and et, a. l. Identification of conserved antigenic components for a cytotoxic T lymphocyte-inducing vaccine against malaria. *Lancet* 345(8956): 1003-7: 1995.

Callan, M.F.C., Steven, N., Krausa, P., Wilson, J.D.K., Moss, P.A.H., Gillespie, G.M., Bell, J.I., Rickinson, A.B., McMichael, A.J. Large clonal expansion of CD^+ T cells in acute infectious mononucleosis. *Nature Medicine.* 2: 906-11: 1996.

Conway, D. J., Holland, M. J., Cambell, A. E., Bailey, R. L., Krausa, P., Peeling, R., Whittle, H. C. and Mabey, D. C. W. HLA class I and II polymorphisms and trachomatous scarring in Chlamydia trachomatis - endemic population. *Journal of Infectious Diseases* 174: 643-646.: 1996.

Imreh, M. P., Zhang, Q. J., Decamposlima, P. O., Imreh, S., Krausa, P., Browning, M., Klein, G. and Masucci, M. G. Mechanisms of allele-selective down-regulation of HLA class-I in burkitts-lymphoma. *International Journal of Cancer* 62(1): 90-96: 1995.